

108.0001

SECTOR \$
PATENT #3

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Jeffrey L. Eakin

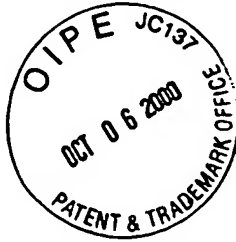
Serial No.: 09/619,411

Filed: July 19, 2000

For: METHODS AND APPARATUS FOR
PROCESSING AND DISTRIBUTING
INFORMATION RELATING TO
COSTS AND SALES OF PRODUCTS

Group: 2753

Examiner: Not Assigned Yet.



I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner of Patents and Trademarks, Washington, D.C. 20231, on the date set forth below:

Signed: Tlynthia Jordan

Name: Tlynthia Jordan

Date: October 4, 2000

Chapel Hill, North Carolina
October 4, 2000

Assistant Commissioner of Patents
and Trademarks
Washington, DC 20231

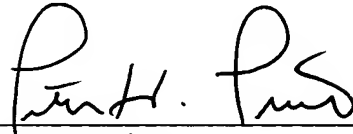
**RESPONSE TO NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL
APPLICATION**

Sir:

Enclosed is a copy of the Notice to File Missing Parts of Nonprovisional Application and Declaration and Power of Attorney for the above identified case.

Also enclosed is our check in the amount of \$130 to cover the fee.

Respectfully submitted,

A handwritten signature in black ink, appearing to read "Peter H. Priest", written over a horizontal line.

Peter H. Priest
Reg. No. 30,210
Law Offices of Peter H. Priest
529 Dogwood Drive
Chapel Hill, NC 27516
(919) 942-1434



UNITED STATES PATENT AND TRADEMARK OFFICE

COMMISSIONER FOR PATENTS
UNITED STATES PATENT AND TRADEMARK OFFICE
WASHINGTON, D.C. 20231
www.uspto.gov

APPLICATION NUMBER	FILING/RECEIPT DATE	FIRST NAMED APPLICANT	ATTORNEY DOCKET NUMBER
09/619,411	07/19/2000	Jeffrey Lynn Eakin	108.0001

Peter H Priest
Law Offices of Peter H Priest
529 Dogwood Drive
Chapel Hill, NC 27516



FORMALITIES LETTER



OC000000005394773

Date Mailed: 09/13/2000

NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL APPLICATION

FILED UNDER 37 CFR 1.53(b)

Filing Date Granted

An application number and filing date have been accorded to this application. The item(s) indicated below, however, are missing. Applicant is given TWO MONTHS from the date of this Notice within which to file all required items and pay any fees required below to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

- The oath or declaration is missing.
A properly signed oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date, is required.
- To avoid abandonment, a late filing fee or oath or declaration surcharge as set forth in 37 CFR 1.16(e) of \$130 for a non-small entity, must be submitted with the missing items identified in this letter.
- The balance due by applicant is \$ 130.

*A copy of this notice **MUST** be returned with the reply.*

Customer Service Center

Initial Patent Examination Division (703) 308-1202

PART 2 - COPY TO BE RETURNED WITH RESPONSE

10/11/2000 WABRHAH1 00000006 09619411

01 FC:105

130.00 OP



UNITED STATES PATENT AND TRADEMARK OFFICE

COMMISSIONER FOR PATENTS
UNITED STATES PATENT AND TRADEMARK OFFICE
WASHINGTON, D.C. 20231
www.uspto.gov

APPLICATION NUMBER	FILING DATE	GRP ART UNIT	FIL FEE REC'D	ATTY. DOCKET NO	DRAWINGS	TOT CLAIMS	IND CLAIMS
09/619,411	07/19/2000	2761	708	108.0001	25	21	3

Peter H Priest
Law Offices of Peter H Priest
529 Dogwood Drive
Chapel Hill, NC 27516



FILING RECEIPT



OC000000005394772

Date Mailed: 09/13/2000

Receipt is acknowledged of this nonprovisional Patent Application. It will be considered in its order and you will be notified as to the results of the examination. Be sure to provide the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION when inquiring about this application. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. If an error is noted on this Filing Receipt, please write to the Office of Initial Patent Examination's Customer Service Center. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the PTO processes the reply to the Notice, the PTO will generate another Filing Receipt incorporating the requested corrections (if appropriate).

Applicant(s)

Jeffrey Lynn Eakin, Residence Not Provided;
Christopher Huyette Howe, Residence Not Provided;
Joseph Eric Lipovich, Residence Not Provided;
Mark Wesley, Residence Not Provided;
Michael James Platt, Residence Not Provided;

mac KENZIE

Please Correct

Continuing Data as Claimed by Applicant

Foreign Applications

If Required, Foreign Filing License Granted 09/12/2000

Title

Methods and apparatus for processing and distributing information relating to cost and sales of products

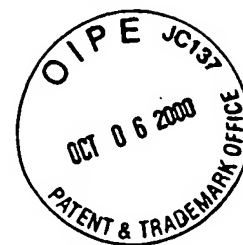
COSTS

Preliminary Class

705

IN THE UNITED STATES
PATENT AND TRADEMARK OFFICE

Declaration and Power of Attorney



As the below named inventors, we hereby declare that:

Our residence, post office address and citizenship are as stated below next to our names.

We believe we are the original, first and joint inventors of the subject matter which is claimed and for which a patent is sought on the invention entitled **METHODS AND APPARATUS FOR PROCESSING AND DISTRIBUTING INFORMATION RELATING TO COSTS AND SALES OF PRODUCTS**, the specification of which was filed on July 19, 2000 as U.S. Serial No. 09/619,411.

We hereby state that we have reviewed and understand the contents of the above identified specification, including the claims, as amended by an amendment, if any, specifically referred to in this oath or declaration.

We acknowledge the duty to disclose all information known to us which is material to patentability as defined in Title 37, Code of Federal Regulations, 1.56.

We hereby claim foreign priority benefits under Title 35, United States Code, 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

None

We hereby claim the benefit under Title 35, United States Code, 119(e) of any United States provisional application(s) listed below:

None

We hereby claim the benefit under Title 35, United States Code, 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, 112, we acknowledge the duty to disclose all information known to us to be material to patentability as defined in Title 37, Code of Federal Regulations, 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

None

We hereby declare that all statements made herein of our own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any

patent issued thereon.

We hereby appoint the following attorney with full power of substitution and revocation, to prosecute said application, to make alterations and amendments therein, to receive the patent, and to transact all business in the Patent and Trademark Office connected therewith:

Peter H. Priest

(Reg. No. 30,210)

Please address all correspondence to Peter H. Priest, Law Offices of Peter H. Priest, 529 Dogwood Drive, Chapel Hill, North Carolina 27516. Telephone calls should be made to Peter H. Priest by dialing Area Code 919-942-1434.

Full name of 1st joint inventor: Jeffry Lynn Eakin

Inventor's signature  Date: 9/25/00

Residence: Ridgefield, CT 06877
Citizenship: US
Post Office Address: 125 Eleven Levels Road
Ridgefield, CT 06877

Full name of 2nd joint inventor: Christopher Huyette Howe

Inventor's signature  Date: 9/25/00

Residence: Stamford, CT 06902
Citizenship: US
Post Office Address: Apt. S-302
1450 Washington Blvd.
Stamford, CT 06902

Full name of 3rd joint inventor: Joseph Eric Lipovich

Inventor's signature  Date: 9/25/00

Residence: Trumbull, CT 06611
Citizenship: US
Post Office Address: 11 Moosehill Rd.
Trumbull, CT 06611

Full name of 4th joint inventor: Mark Wesley Mac Kenzie

Inventor's signature Mark Wesley MacKenzie Date: 9/25/00

Residence: Brewster, NY 10509
Citizenship: US
Post Office Address: 81 Acorn Road
Brewster, NY 10509

Full name of 5th joint inventor: Michael James Platt

Inventor's signature Michael James Platt Date: 9/25/00

Residence: Fairfield, CT 06432
Citizenship: US
Post Office Address: 279 Lynnbrook Rd.
Fairfield, CT 06432

Position	Sequence	Position	Sequence	Position	Sequence
Db	2548 GGCTGGGGGGGTGGCTCACCCTTAATATCCACGATTTTGGAGGCCGAGGGGTGAA	2607			
QY	805 TCACCTTAAGSTCAGAGATTGTAGACCCAGCCTTGCCACATGTGTGTAAGGTGTCTCACT	864			
Db	2608 TCACCTTGAGGTCAGAGATTTCAGAGCCACGACCTGGCCACATGTGTGTAAGCTTGTCTCACT	2667			
QY	865 AAAAATACAAAAATTAGACAGGCGTGTGGGACACATCTGTATTTCCAGTACTCAGAG	924			
Db	2668 AAAAATACAAAAATTAGACAGGAGTGTGGGACACCTGTATTTCTGTACTACGAGAG	2727			
QY	925 GCTAACACAGGAAAATTCCTTGAACCTGTGGAGGACAGAGTTGACGTGACCATTCACCTC	984			
Db	2728 GCTGAGTAGAGAGAACTGCTTGCACCCAGGAGGACAGAGTTGCAATGTGGCCACTGCACTC	2787			
QY	985 CAGCTCTGGGCAACACAGTGAAGCTTGTCTTCAAAAAA	1030			
Db	2788 CAGCTCTGGTGCACAAAGGAGACACTGTGCACACACATATACACAA	2833			

Accession	Sequence	Position
Qy	GGTGGGCGCCAAATATATGCTCTGGGCGCAAGCTTTTCGGGGCGAGAGGCAAGTGGAA	80
745	GGTGGGCGCCAAATATATGCTCTGGGCGCAAGCTTTTCGGGGCGAGAGGCAAGTGGAA	80
308	GGCTGGGATATGGTGGCTCAAGCTCTTAATCCAGCATTTTGGAGGCCAAGCAGCTGGA	249
Db		
805	TCACTTAAGTCAGAGATTGTAGACACAGCTGGCCCAACATGTGTAAGACGTTGCTCTACT	86
Yy		

Db	248	TACCTGAGGTGAGGAGTTTGAGACAGCCTGGCCAAATATGGCTTAAACCCCATCTTACT	189		
Qy	865	AAAAATACAAAATTAAGACAGCGCGTGGTGACACATCTGTATTCACAGTACTACGAG	924		
Db	188	AAAAATACAAAATTAAGACAGCGCGTGGTGACAGCCTGTATTCACAGTACTTGGAG	129		
Qy	925	GCTAACACAGGAAATATTCCTTGAACCTGGGAGGCAAGGTTGACGTGAG-----	973		
Db	128	GCTGAGGACGAGAAATGCTTGATCTCTGGAGGCAAGGTTGACGTGAGTACCA	69		
Qy	974	CCATTGCACTCCAGCCTGGGCAACACAGTGAGCTTGTCTCAAAAAA	103		
Db	68	CCATTGCACTCCAGCCTGGGCAACAGATGAGCTTGTCTCAAAAAA	9		
RESULT 2	BM907568/c	1024 bp	mRNA	linear	EST 12-MAR-2000
LOCUS	BM907568				
DEFINITION	AGENCOURT 6641246 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:5735872				
ACCESSION	5', mRNA sequence.				
VERSION	BM907568				
KEYWORDS	BM907568.1 GI:19357947				
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
JOURNAL	1 (bases 1 to 1024)				
COMMENT	NIH-MGC http://mgs.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/DRP/gazdar cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: L14M12743 row: h column: 17 High quality sequence stop: 518. location/qualifiers 1..1024 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5735872" /issue_type="large cell carcinoma" /lab_host="DPH10B (phage-resistant)" /clone_id="NIH_MGC_68" /note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."				
FEATURES					
SOURCE					
ORIGIN					
Query Match	18.9%;	Score 196.6;	DB 5;	Length 1024;	
Best Local Similarity	66.2%;	Pred. No. 9.2e-22;			
Matches 304;	Conservative 0;	Mismatches 144;	Indels 11;	Gaps 1;	
Qy	586	AATTTTATTAATATTAAGAGTGAAGTATGACATGACATGACAAACAGGAAATATAA	645		
Db	461	AAAGTATATTAATTAATTAATTAAGTAACTAAGCTGAGATGTTACACCAAGTAAGAAAGCA	402		
Qy	646	CAGATTTCCGTTTGCGCTTGTGCTAAATTTGTAACCTGACAAATCTTACCGAGTTCTTCA	705		
Db	401	CAGGTAAATAGGCTTCTCTCCCTAATACCAAGTAATTTATTAACCTACACAGATTGGCC	342		
Qy	706	TTTTGCTTTGAGATGAGATTAATTAAGTTTAACTGCAAGGGGCTGGGCGGATAGCTGATG	765		
Db	341	AATTTAGTAAATTAATTAATTAATTAAGTAAATTTTAAAGTATTTTAAGGCGGCGATGTTGGCTACA	282		

QY	766	CCTTGGGTCGAGNCTTTGCGGGGGCCGAGCAGGTGCATCTTAAGTCAGAGACTTTG	8235	
Db	281	CTTATATCCGACGACTTTTGGAGGGCCGAGCGGGCGGATCACCTGAAGTCAGAGACTTTG	2222	
QY	826	AGACACGACTGCCCAACATGATGTAACCTTTGTCTCTACTAATAAAATACAAAATAATTAGACAG	8855	
Db	221	AGACACACCTGCAGCAACATGATGTAACCTGCATTTCTAATAAAATAAAATAATTAGACAG	1622	
QY	886	GGCGGGGGGACACATCTGTAAATTTCCAGCTACTCAGAGAGCTAACACAGAAATTCCTT	9455	
Db	161	GTGTGTGGATGTGCTCTGTAATCCAGCTACTCAGAGAGCTGAGACAGAGAAATTCCTT	1022	
QY	946	GAACTTGGAGGAGGAGTTCAGTGA-----GCCATTGCACCTCGAGCTCGGG	9944	
Db	101	GAACCTGGGAAGCAGACGCTGCAGTACGATGATGTCGATTCGATTCGAGCTCGGG	4222	
QY	995	AACACAGTACGACTCTTTGTCTCAAAAAAAAAAAAAAAAAAAAAA	1033	
Db	41	AACACAGAGCGAAATTCGCTCTCAAAAAAAAAAAAAAAAAAAAAA	3333	
RESULT 3	AI733856	458 bp	linear	EST 24-OCT-2000
LOCUS	2019c03.y5	Stratagene colon (#937204)	Homo sapiens	CDNA clone
DEFINITION	IMAGE:587332.5'	similar to contains Alu repetitive element;	mRNA	sequence.
ACCESSION	AI733856			
VERSION	AI733856.1	GI:5054969		
KEYWORDS	EST.			
SOURCE	Homo sapiens	(human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 458)			
AUTHORS	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index			
JOURNAL	Unpublished (1997)			
COMMENT	Other_ESTs: 2019c03.x5			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgapbs-remail.nih.gov			
	This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.			
	This read is a RESSEQUENCE of a previously sequenced human clone			
	Original clone citation: Washu-Merck EST Project			
	This read has been verified (found to hit its original self in the correct orientation)			
	Insert Length: 1908	Std Error: 0.00		
	Seq primer: -40RP from Gibco			
	High quality sequence stop: 381.			
FEATURES	location/Qualifiers			
source	1..458			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="GDB:4619634"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:587332"			
	/tissue="tumor"			
	/cell_line="T84 carcinoma cell line"			
	/lab_host="SOLR cells (kanamycin resistant)"			
	/clone_id="Stratagene colon (#937204)"			
	/note="Organ: colon; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cloned unidirectionally. Primer: Oligo dt. T-84 colonic epithelial cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGACGAG 3' ~3' adaptor sequence: 5' CTCGAGCTTTT TTTT TTTT 3' "			
ORIGIN				
Query Match	18.9%	Score 196.4	DB 1	Length 458
Best Local Similarity	76.3%	Prod. No. 1.2e-21		

Matches 257; Conservative 0; Mismatches 66; Indels 11; Gaps 1;

QY 716 GAGATGAAGTACTTACTTATTTAGTCCAGGGGCTGGCCGATAGCTATGCTTGTGCTCC 775
DB 346 GAATTAATAATATATCTTTTATTTTGGCCGAGGGGTGTGCTCATGCTGTATATCC 287
QY 776 CAGTGTCTTGGGGGCGGAGGAGTGAATCACTTAAGTTCAGAGTTTGGAGACAGCCT 835
DB 286 CAGACCTTTGGAGTCCGAGGAGAGTGAATCACTTGAGTTCAGAGTTCGAGACCACTT 227

QY 836 GCCCAACATGATGTAAGCTTGTCTCTACTAATAAATAACAAAATTAGA CAGGCGTGTGCG 895
DB 226 GACCAACATGATGTAAGCTTGTCTCTAATAAATAACAAAATTAGA CAGGCGTGTGCG 167

QY 896 ACACATCTGTAATTCAGCTACTCAGAGGCTTAACAAGAAATTCCTTGAACCTGGGA 955
DB 166 ATGGCCCTTAATATCCAGCTACTCAGAGGCGTGAAGAGAGAAATCCCTTGAACCTGGGA 107

QY 956 GGCAGAGTTCGAGTGA-----GCCATTGCACTCGAGCCTGGGCAACAGTGA 1004
DB 106 GGTGAGAGTTCGAGTGAAGCCGAGATCGTCCACTGCACTCGAGCCTGGGAGAGCA 47

QY 1005 GACTCTTGTCTCAAAAAAAAAAAAAAACTCGA 1038
DB 46 GACTCCATCTCAAAAAAAAAAAAAAACTCGA 13

RESULT 4
BO004374/c 621 bp mRNA linear EST 26-MAR-2002
LOCUS UI-H-E10-849-j-13-0-UI.s1 NCI CGAP_E10 Homo sapiens cDNA clone
DEFINITION IMAGE:5841132 3', mRNA sequence.
ACCESSION BO004374
VERSION BO004374.1 GI:19729274
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 621)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@ds-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
The following repetitive elements were found in this cDNA
sequence: 12-301, >ALU (matched complement)
Seq primer: M13 FORWARD
POLYA=yes

FEATURES
source Location/Qualifiers
1. 621
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5841132"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="NCI CGAP E10"
/note="Organ: Left Pelvis; Vector: p773-Pac (Pharmacia)
with a modified polylinker. Site 1: Ecor 1; Site 2: Not I;
NCI-CCAP_E10 is a cDNA library containing the following
tissue(s): Chondrosarcoma. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into p773-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is ACACCTGCAC.
TAG TISSUE=chondrosarcoma
TAG LIB=UI-H-E10
TAG_SEQ=ACACCTGCAC"

ORIGIN

Query Match 18.9%; Score 196; DB 5; Length 621;
Best Local Similarity 81.3%; Pred. No. 1.3e-21;
Matches 244; Conservative 0; Mismatches 45; Indels 11; Gaps 1;

QY 745 GCGTGGCGCGATGCTGATGCTGTGCTCCAGTCTTTGGGGCGGAGCGAGTGA 804
DB 301 GCGTGGGTGTGTGCTGCTCAGCGCTGTATCTCAGCACTTTGGGAGCGGAGCGAGTGA 242

QY 805 TCACCTTAAGGTCAGGAGTTCAGACCGAGCCTGCCCAATGCTGAACGTTGTCTACT 864
DB 241 TCATCTAGGTTCAGAGATTCGAGACAGCTGAGCAACATGTAACCTGTCTGCT 182

QY 865 AAAAATCAAAAATTAGACAGGCGGTGGGACACATCTGTATTCAGTACTCAGAG 924
DB 181 AAAAATCAAAAATTAGCCAGGTGTGGGAGCGGCTGTATTCAGTACTCAGAG 122

QY 925 GCTAACACAGAAAATTCCTTGAACCTGGAGGAGAGGTTGACGTGA-----G 973
DB 121 GCTGAGGACAGAGAAATCCCTTGAAGCTGAGGAGGAGGTTGACGTGAAGCAGATCGTG 62

QY 974 CCAATTGACCTCCAGCCTGGGCAACAGAGTCTGTCTCAAAAAAAAAAAAAA 1033
DB 61 CCAATTGACCTCCAGCCTGGGATAGAGTCTGTCTCAAAAAAAAAAAAAA 2

RESULT 5
BO017611/c 554 bp mRNA linear EST 17-JUN-2002
LOCUS UI-H-D10-auv-j-09-0-UI.s1 NCI CGAP_D10 Homo sapiens cDNA clone
DEFINITION IMAGE:5874920 3', mRNA sequence.
ACCESSION BO017611
VERSION BO017611.1 GI:19752888
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 554)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@ds-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
The following repetitive elements were found in this cDNA
sequence: 1-27, >AT-richlow_complexity 133-417, >ALU (matched
complement)
Seq primer: M13 FORWARD
POLYA=yes

FEATURES
source Location/Qualifiers
1. 554
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5874920"
/tissue_type="Lung Focal Fibrosis"

```
/dev stage="Adult"  
/lab host="DH10B (Life Technologies)"  
/clone lib="NCI_CGAP_D10"  
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
NCI_CGAP_D10 is a cDNA library containing the following  
tissue(s): A pool of lung focal fibrosis. The library was  
constructed according to Bonaldo, Lemon and Soares,  
Genome Research, 6:791-806, 1996. First strand cDNA  
synthesis was primed with an oligo-dT primer containing a  
Not I site. Double stranded cDNA was ligated to an EcoR I  
adaptor, digested with Not I, and cloned directionally  
into pT73-Pac vector. The oligonucleotide used to prime  
the synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(dT)18 tail. The sequence tag for this library is  
ATACGCGTC.  
TAG_TISSUE=Lung with fibrosis  
TAG_LIB=UI-H-D10  
TAG_SEQ=ATACGCGTC"
```

ORIGIN

```
Query Match 18.9%; Score 195.8; DB 5; Length 554;  
Best Local Similarity 78.7%; Pred. No. 1.5e-21;  
Matches 251; Conservative 0; Mismatches 57; Indels 11; Gaps 1;
```

```
731 AGTTTATGTCAGGAGCTGGGCGCATAGCTGATGCTGTGGTCCAGTGGCTTTGGCGGG 790  
431 AATATGTGTATAGCTGGGCGGTGGCTCATCTGTATCCACACCTTTGGGGGG 372  
731 CCGAGGCGAGGTGATCACTTAAGTCAAGTTTGAACCAAGCTGCCCAATGTGAA 850  
371 CCAGGCGAGGTGATCACTGAGTCAAGATTGAACCAAGCTGCCCAATGTGAA 312  
851 AGTGTGCTCTACTTAAATCAAAATTTAGACGGCGGTGGGACACATCTTAATTC 910  
311 ATCCAGTCTCTTAAATCAAAATTTAGACGGCGGTGGGACATTTGTATATCC 252  
911 CAGTACTCAGAGGCTTAACACAGGAAATTCCTTGAACCTGGAGGACAGAGTTGCAGT 970  
251 CAGTACACACAGAGGCTGAGGACGAAATTCCTTGAACCTGGAGGAGGAGTTGCAGT 192  
971 GA-----GCCATGCACTCCAGCTGGGCAACAGTGAAGCTTGTCTCAA 1019  
191 GAACCAAGTGGCGCATGCACTCAGCTGGGCGACAGAGTGAACCTTTCAAAAAC 132  
1020 AAAAAAAAAAAAACTGCA 1038  
131 AAAAAAAAAAAAAAGA 113
```

```
RESULT 6  
T78471/c 559 bp mRNA linear EST 15-MAR-1995  
LOCUS yd68f03.81 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
DEFINITION IMAGE:113405 3' similar to contains Alu repetitive element; mRNA  
sequence.  
T78471  
T78471.1 GI:696980
```

```
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

```
REFERENCE  
AUTHORS  
1 (bases 1 to 559)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., and  
Wilson, R.
```

```
TITLE  
JOURNAL  
COMMENT  
The WashU-Merck EST Project  
Unpublished (1995)  
Other_ESTs: yd68f03.r1
```

```
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
Insert Size: 632  
High quality sequence stops: 489 Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 632 Std Error: 0.00  
Seq primer: -21m13  
High quality sequence stop: 489.  
Location/Qualifiers  
1. 559  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="GDB:469022"  
/db_xref="taxon:9606"  
/clone="IMAGE:113405"  
/sex="male"  
/dev stage="20 week-post conception fetus"  
/lab host="DH10B (ampicillin resistant)"  
/clone lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACTGGAAGATTAATTAAGATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT  
3']  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pT73 vector. Library  
constructed by Benito Soares and M.Patima Bonaldo."
```

ORIGIN

```
Query Match 18.8%; Score 195.6; DB 7; Length 559;  
Best Local Similarity 72.1%; Pred. No. 1.6e-21;  
Matches 266; Conservative 0; Mismatches 102; Indels 1; Gaps 1;
```

```
665 TTGCTAAATTTGTACTGACAAATCTTAAGCCAGTTCTTCAATTTGGTTGAGATGAG 724  
370 TGTATAATATTTTGAACAGGAAATTAAGTCTGTCTTCTTGGTTTGAAGGCTTAAAG 311  
725 ATACTTAATTTTGAACAGGAGCTGGGCGCATAGCTGATGCTGTGATCCAGGCTTT 784  
310 TACATATTAGAAATTTTGAAGCTGGGCGCATAGCTGATGCTGTGATCCAGGCTTT 251  
785 GCGGGGCGAGGCGAGGTGATCACTTAAGTCAAGAGTTTGAACCAAGCTGCCCAAT 844  
250 GTGAGGCGGACGACAGGTGATCACTTGAAGTCA -GAATTTGAGACCAAGCTGGCCACAC 192  
845 GGTGAAGGTTGTCTCTCTAATAAATTAACAAATTAAGACAGCGGTGGGACACATCTG 904  
191 AGTTAAGCCCACTCTTCTAATAAATTAACAAATTTGGCCAGCAAGTGGGGGACCTG 132  
905 TAATTCAGCTTACTCAGAGGCTTAACACAGGAAATTCCTTGAACCTGGAGGAGAGGT 964  
131 TAATCCAGCTACTTGGAGAGATGAGGAGGAATCGCTTGAACCTGGAGGAGAGGT 72  
965 TGCATGAGCCATTTGCACTCCAGCTGGGCAACACAGTGAAGCTCTTGTCTCAAAAAA 1024  
71 TGCAGTGAAGCATTTGCACTCCAGCTGGGCAACAGAGTGAAGCTCTTGTCAAAAAA 12  
1025 AAAAAAAAA 1033  
11 CAAAAAAA 3
```

```
RESULT 7  
A0527649 456 bp DNA linear GSS 18-MAY-1999  
LOCUS R0527649  
DEFINITION R0527649, genomic survey sequence.  
R0527649, genomic survey sequence.
```


ACCESSION	AF052764.9
VERSION	AF052764.9.1
KEYWORDS	GI:4839803
SOURCE	GSS.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE	1 (bases 1 to 456)
AUTHORS	Zhao, S., Adams, W.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.
TITLE	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL	Unpublished (1997)
COMMENT	Other GSSs: RPCI-11-348021.TU Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbe@tigr.org Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Peter de Jong (pj@redwing.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet. es (info@resgen.com). BAC end search page: http://www.tigr.org/ttdb/hungen/bac_end_search/bac_end_search.html. Seq primer: 77 Class: BAC ends.
FEATURES	Location/Qualifiers
SOURCE	1..456
	/organism="Homo sapiens"
	/mol_type="Genomic DNA"
	/db_xref="GDB:7631604"
	/db_xref="taxon:9606"
	/clone="RPCI-11-348021"
	/sex="Male"
	/cell_type="Lymphocytes"
	/clone_11b="RPCI-11"
	/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; RPCI11 Human Male BAC Library"
ORIGIN	
Query Match	18.8%; Score 195; DB 8; Length 456;
Best Local Similarity	78.0%; Pred. No. 2.1e-21;
Matches 252; Conservative	0; Mismatches 60; Indels 11; Gaps 1
QY	721 GAAGATCTTACTTTAGTCCAGGGGCTGGGGCCATAGCTGATCCTGTGTCCAGTG 780
DB	59 GACCTTAAAAATGTTAAATAGAGCGCTGGGCGATGTGGCTCACACCTGTATCCACAGA 118
QY	781 CTTGGCGGGCGAGGAGGTGATCATTAAAGTCAGAGATTGAGACCGCCTGGCCA 840
DB	119 CTTGGGAGGCGAGGCGGTGATCAGCTGAGTTCAGAGAGTTTCAGACCGCCTGGCCA 178
QY	841 ACATGCTGAACGTTGCTCTACTTAAATACAAAAATTGACAGCGCTGTGTGACACA 900
DB	179 ACATGGCGAACCCTGCTCCACTTAAATACAAAAATTAGTCAGGCGTGTGGCATGTG 238
QY	901 TCTGTAATTCAGCTACTCAGAGGCTTAACACGAAAAATCTTGAACCTGGAGGCA 960
DB	239 TCTGTAATTCAGCTACTCAGAGGCTTAACACGAAAAATCTTGAACCTGGAGGCA 298
QY	961 AGGTTGACGTAG-----CCATTGCACTCAGCGTGGGCAACACGTGAGACTC 1000
DB	299 AAGTTGACGTAGCCGAGATGCACACTGTCAGCTCAGCTGGGTAAACAAACGAGACTC 358
QY	1010 TTGTTCTCAAAAAAAAAAAAAAA 1032
DB	359 CTGTCTCAAAAAAAAAATCTTAAAA 381
RESULT 8	

LOCUS	AC0481653	551 bp	DNA	linear	GSS 24-APR-1999
DEFINITION	RPCT-11-235F3.TV RPCT-11 Homo sapiens genomic clone RPCT-11-235F3, genomic survey sequence.				
ACCESSION	AC0481653				
KEYWORDS	AC0481653.1 GI:469057				
SOURCE	GSS.				
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 551) Zhao,S., Adams,M.D., Niernan,W., Matek,J., de Jong,P. and Venter,J.C. End Sequences from Library RPCT-11 for Sequence-Ready Use of BAC End Sequences from Library RPCT-11 for Sequence-Ready Map Building Unpublished (1997)				
JOURNAL	Other GSSs: RPCT-11-235F3.TV Contact: Shaying Zhao, William Niernan, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208				
COMMENT	Email: hbe@igr.org Clones are derived from the human BAC library RPCT-11. For BAC library availability, please contact Pieter de Jong (pieter@jeng.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.ctigr.org/ctdb/humgen/bac_end_search/bac_end_search.html. Seq primer: T7 Class: BAC ends.				
FEATURES	Location/Qualifiers				
source	1..551 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="GDB:7589978" /db_xref="taxon:9606" /clone="RPCT-11-235F3" /sex="Male" /cell_type="lymphocytes" /clone_id="RPCT-11" /note="Vector: pBAC3.6; site_1: EcoRI; site_2: EcoRI; RPCT11 Human Male BAC library"				
ORIGIN	Query Match 18.8%; Score 194.8; DB 8; Length 551; Best Local Similarity 79.5%; Pred. No. 2.1e-21; Matches 229; Conservative 0; Mismatches 59; Indels 0; Gaps 0;				
QY	745	GGCTGGGGCGCATGTGATGCTCGTGTGATGCCACGATGCTTTGCGGGGCGAGGACGATGGA	804		
DB	247	GCGTGGGGGTGTGTGGCCACGCGCTGTATATCCACACACTTTTGGAGGCTGAGATAGTGA	306		
QY	805	TCACCTAAGGTACAGAGTTTGAGACCAAGCCCTGCCCAATGATGAAAAGTTGTCTTACT	864		
DB	307	TCACCTGAGGTCAAGGACATTGAGACCAAGCTGACCAACAAGGTAAACCCGTGTCTTACT	366		
QY	865	AAAAATACAAAATTATTAACACAGGCGTGTGGTGCACACATCTGTAAATCCAGCTACTCAGAG	924		
DB	367	AAAAATACAAAATTATTAACCTGGGCGTGTGGAGGCTCTTGTAGTCCACAGCTACTCGGGGG	426		
QY	925	GCTAACACAGGAAATTCCTTGAACCTTGGAGGAGGAGAGGTTGACGATGAGCCATTGCACTC	984		
DB	427	GCTAGAGAGGAGGAATTTGCTTTGAACCCGGAGGCGAGAGGTTGCATGAGCCACTGCCTACTC	486		
QY	985	CAGCCTGGGCAACACAGTGAAGACTCTTGTCTCAAAAAAAAAAAAAAAAAAAAAA	1032		
DB	487	CAGTCTGGGTGATGAGTGAAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAA	534		

ACCESSION BM67682
 VERSION BM67682.1 GI:18975513
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 504)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA sequence: 1-27, >AT_richlow_complexity 132-416, >Alu (matched complement)
 Seq primer: M13 Forward
 PolyA=yes.

FEATURES
 source Location/Qualifiers
 1..504
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-DX0-agg-f-14-0-UI"
 /tissue="fetal eyes"
 /dev_stage="fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_1lb="UI-E-DX0"
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-DX0 is a cDNA library containing the following tissue(s): fetal eyes. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGAATCAAGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
 TAG_TISSUE=human fetal eyes
 TAG_LIB=UI-E-DX0
 TAG_SEQ=AGAATCAAGA"

ORIGIN
 Query Match 18.7%; Score 194.2; DB 4; Length 504;
 Best Local Similarity 78.4%; Pred. No. 2.7e-21;
 Matches 250; Conservative 0; Mismatches 58; Indels 11; Gaps 1;

Gy 731 AGTTTATGTCAGAGGCTGGCGCGATAGCTGATGCTGCTCCAGTCTTGGCGGG 790
 |||
 Db 430 AATATGTGTGTATAGGCTGGGTGGTGGCTGATGCTGTATATCCAGACATTTGGAGG 371
 |||
 Gy 791 CCGAGGAGGTGATCACTTAAGTCAAGAGCTTTGAACCGAGCTGCCCAACATGTGAA 850
 |||
 Db 370 CCAAGGAGGTGATCACTTGAAGTCAAGAGTTTGAAGCCAGCTGACCAACATGTGAA 311

RESULT 12
 A1073373/c 515 bp mRNA linear EST 06-AUG-1998
 LOCUS A1073373/c
 DEFINITION ool3a11.x1 Soares NSF Pg. 9M OT PA P S1 Homo sapiens cDNA clone IMAGE:1566044 3' similar to contains Alu repetitive element; mRNA sequence.
 ACCESSION A1073373
 VERSION A1073373.1 GI:3400017
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 515)
 REFERENCES NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: rga@pfeil.nih.gov
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 454.
 Location/Qualifiers
 1..515
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1566044"
 /lab_host="DH10B"
 /clone_1lb="Soares NSF Pg. 9M OT PA P S1"
 /note="Organ: pooled; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and as circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and clones: Soares NBHSF pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152337 Soares NB2HP-9M pool 1: 758280-760583, 772104-774407 Soares NBHRA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NBHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 18.7%; Score 194.2; DB 1; Length 515;
 Best Local Similarity 78.4%; Pred. No. 2.7e-21;
 Matches 250; Conservative 0; Mismatches 58; Indels 11; Gaps 1;

Gy 731 AGTTTATGTCAGAGGCTGGCGCGATAGCTGATGCTGCTCCAGTCTTGGCGGG 790

Db	423	AATATGTGTATAGAGCTGGGTGGGTGACTATGCTTATCCAGCATTTGGGAGG	364
QY	791	CCGAGGCAAGTGGATCATTTAAAGTCAAGGATTTGGAGACCAAGCTGGCCCAACATGTGTAA	850
Db	363	CCAAAGCAGTGGATCATCCTTGAGGTCAAGGAGTTTGAGACCAAGCTTACCAACATGTTGAA	304
QY	851	ACGTTGTCTCTACTTAAAAATACAAAAATTAGACAGGCGTGGTGGCACACATCTGTATATTC	910
Db	303	ATCCAGTCTCTACTTAAATATACAAAAATTATGACAGGCGTGGTGGCTGTGATTTGTATATCC	244
QY	911	CAGCTACTCAGGAGGCTAACACAGGAAAAATTCCTTGAACCTGGAGGCGAGAGTTCAGGT	970
Db	243	CAGCTAACACAGGAGGCTGAGCGAGAGAAATTCCTTGAACCTTGGGGGGCGGAGGTTTGCAT	184
QY	971	GA-----GCCATTGCACCTCAGGCTGGGCAACACAGTGAAGACTTGTGTCTCAA	1015
Db	183	GAGCCAAAGTGGCGGCATTTGCACCTCAGGCTGGGCGACAGAGTGAACATTCTCAAAAAAC	124
QY	1020	AAAAAAAAAAAAAACTGCA	1038
Db	123	AAAAAAAAAAAAAAACA	105

RESULT 13	LOCUS	DEFINITION	ACCESSION
A0314507/c	A0314507	524 bp DNA linear	GSS 04-MAY-1999
	RPCT11-103B6	TV RPCT-11 Homo sapiens genomic clone	RPCT-11-103B6,
		genomic survey sequence.	
			U014507

ACCESSION AQ314507
 VERSION AQ314507.1 GI:4045970
 KEYWORDS GSS.

SOURCE ORGANISM	Homo sapiens (human)
	Homo sapiens

REFERENCE
1 (bases 1 to 524)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE	AUTHORS
Use of human BAC End Sequences for Semiautomated Mapping of Human DNA	Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P., and Venter, J.C.

JOURNAL
COMMENT
Unpublished (1998)
Other_GSSs: RPC111-103B6.TJ
Contact: Chandra Bhargava, 1111

contact: suaying zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208

Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, see <http://www.tigr.org>

For further availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from

Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: T7

```

Class: BAC ends.
FEATURES
source      1      524

```

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
```

```
/db_xref="GDB:7539197"  
/db_xref="taxon:9606"  
/clone="RPC1-11-103B6"
```

```
/sex="Male"  
/cell_type="Lymphocytes"  
/clone_id="PBCT-11"
```

Recoil Human Male BAC Library"

ORIGIN	Query Match	18.7%	Score 194.2	DB 8	Length 524
--------	-------------	-------	-------------	------	------------

Best Local Similarity 79.8%; Pred. No. 2.7e-21;

	Matches	229;	Conservative	0;	Mismatches	58;	Indels	0;	Gaps	0;
QY	743	GGGGGTGGGGCCCGAATAGCTGATGCTGCTGGTGTCCCAAGTCTTTGCGGGGCCGAGGCGAGGTG								802
Db	415	GAGGCTGGGCGCGGGTGGCTCAATGCTGTAATTCAGACATTGGAGAGGCCAGAGTGGGCG								356
QY	803	GATCATTTAAGTCAGAGATTGAGACCCAGCCTGCCCAACATGGTGAACGTTGCTCTTA								862
Db	355	GATCACCTGAGAGTCAGGAGTTGAGACCCAGCCTGAGCCACATGGTGAACCCCACTCTTA								296
QY	863	CTAAAAATATCAAAAATTTAGACAGGCGTGGTGGACACATCTGTAATTCACACTACTAG								922
Db	295	CTAAAAATATCAAAAATTTAGCCAGGCGTGGTGGCATGTGCTGTAAATCCAGACTACTAG								236
QY	923	AGGTTAAACACAGGAAAATTCCTTTGAACCTGGGAGGCGAGAGTTGCACTGAGCCATTGCAC								982
Db	235	AGGCTTAAAGGCGCGGAATTCGCTTGAATCTCAAGAGGCGAAGATTGCACTGAAACCAACGCGCAC								176
QY	983	TCCAGCCTGGGCAACACAGTGAAGACTTGTCTCTCAAAAAAAAAAAAAA								1029
Db	175	TCCAGCCTGGGTTACGAGCGAAGTCTATCTCTCAAAAAAAAAAAAAA								129

Accession	EU626395	724 bp	linear	EST 23-SEP-2002
Definition	UI-H-DFO-beu-o-10-0-UI s1 NC1 CGAP DFO Homo sapiens cDNA clone			
LOCUS	UI-H-DFO-beu-o-10-0-UI 3', mRNA sequence.			
Accession	EU626395			

VERSION	BU626395.1	GI:23292610
KEYWORDS	EST.	
SOURCE	HOMO sapiens (human)	

ORGANISM Homo sapiens (human)
Eukaryota: Metazoa:

REFERENCE	REFERENCE
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo.

FOUNDRING
TITLE
NCI-CCGAP <http://www.ncdi.nlm.nih.gov/ccgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: c9abbs-r@mail.nih.gov

Tissue Procurement: Dr. José Mercuende
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA library Arranged by: Dr. M. Bento Soares, University of Iowa

Cloned and sequenced by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained

The following repetitive elements were found in this cDNA sequence: 1-27, >AT rich#Low complexity (matched complement from Dr. M. Bento Soares, bento-soares@uiowa.edu)

133-417, >ALU (matched complement)
Seq primer: M13 FORWARD
POLYA-Yes

FEATURES	Location/Qualifiers
source	1. .724

```
/organism="Homo sapiens"  
/mol_type="mrna"  
/db_xref="taxon:9606"
```

```
/clone="UI-H-DF0-beu-o-10-0-UI
/tissue_type="Subchondral Bone
/dev stage="Adult"
```

```
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_DF0"
```

/note="Organ: Bone; Vector: pLV3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I, NCI CGAP DF0 is a cDNA library containing the following

tissue(s): Subchondral Bone. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested

vector. The oligonucleotide used to prime the synthesis of

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 07:53:15 ; Search time 59 Seconds
(without alignments)
61.997 Million cell updates/sec

Title: US-10-664-358-549

Perfect score: 260

Sequence: 1 MNLIGMIFSMCGMLKLMKWC.....ISFANSRSSEDTKOMSSFM 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:*

2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*

7: /cgn2_6/ptodata/1/1aa/6D.COMB.pep:*

8: /cgn2_6/ptodata/1/1aa/6E.COMB.pep:*

9: /cgn2_6/ptodata/1/1aa/6F.COMB.pep:*

10: /cgn2_6/ptodata/1/1aa/6G.COMB.pep:*

11: /cgn2_6/ptodata/1/1aa/6H.COMB.pep:*

12: /cgn2_6/ptodata/1/1aa/6I.COMB.pep:*

13: /cgn2_6/ptodata/1/1aa/6J.COMB.pep:*

14: /cgn2_6/ptodata/1/1aa/6K.COMB.pep:*

15: /cgn2_6/ptodata/1/1aa/6L.COMB.pep:*

16: /cgn2_6/ptodata/1/1aa/6M.COMB.pep:*

17: /cgn2_6/ptodata/1/1aa/6N.COMB.pep:*

18: /cgn2_6/ptodata/1/1aa/6O.COMB.pep:*

19: /cgn2_6/ptodata/1/1aa/6P.COMB.pep:*

20: /cgn2_6/ptodata/1/1aa/6Q.COMB.pep:*

21: /cgn2_6/ptodata/1/1aa/6R.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	100.0	49	US-09-482-273-180	Sequence 180, App
2	260	100.0	70	US-09-188-930-313	Sequence 313, App
3	260	100.0	70	US-09-312-283C-313	Sequence 313, App
4	260	100.0	71	US-09-482-273-116	Sequence 116, App
5	260	100.0	106	US-09-247-155-90	Sequence 90, App1
6	260	100.0	106	US-09-513-999C-8115	Sequence 8115, App
7	217	83.5	114	US-09-270-767-33341	Sequence 33341, A
8	217	83.5	114	US-09-270-767-48558	Sequence 48558, A
9	61	23.5	240	US-09-270-767-39637	Sequence 39637, A
10	61	23.5	240	US-09-270-767-54854	Sequence 54854, A
11	58.5	22.5	322	US-09-254-227A-5	Sequence 5, App11
12	54.5	21.0	322	US-09-254-227A-3	Sequence 16, App1
13	53	20.4	65	US-09-465-901-16	Sequence 16, App1
14	52	20.0	65	US-09-248-796A-25087	Sequence 25087, A
15	52	20.0	190	US-09-248-796A-15809	Sequence 15809, A
16	51	19.6	223	US-09-465-901-14	Sequence 14, App1
17	51	19.6	223	US-09-465-901-32	Sequence 32, App1
18	50.5	19.4	316	US-09-252-991A-17312	Sequence 17312, A
19	50.5	19.4	332	US-09-134-000C-5338	Sequence 5338, App
20	50	19.2	139	US-09-270-767-37702	Sequence 37702, A
21	50	19.2	139	US-09-270-767-52919	Sequence 52919, A
22	50	19.2	239	US-09-488-039A-10233	Sequence 10233, A
23	50	19.2	848	US-09-543-681A-7615	Sequence 7615, App
24	49.5	19.0	202	US-09-252-991A-19646	Sequence 19646, A
25	49.5	19.0	585	US-09-248-796A-20096	Sequence 20096, A
26	49.5	19.0	741	US-09-436-699C-22	Sequence 22, App1
27	49	18.8	54	US-09-621-976-5390	Sequence 5390, App

28	49	18.8	118	4	US-09-270-767-62231	Sequence 62231, A
29	49	18.8	154	4	US-09-248-796A-20059	Sequence 20059, A
30	49	18.8	223	4	US-09-465-901-36	Sequence 36, App1
31	49	18.8	348	4	US-09-270-767-46628	Sequence 46628, A
32	49	18.8	373	4	US-09-248-796A-17973	Sequence 17973, A
33	49	18.8	462	3	US-09-165-241-1	Sequence 1, App11
34	49	18.5	125	4	US-09-621-976-5389	Sequence 5389, App
35	48	18.5	125	4	US-08-469-260A-412	Sequence 412, App
36	48	18.5	125	4	US-08-469-260A-412	Sequence 412, App
37	48	18.5	125	4	US-08-467-344A-412	Sequence 412, App
38	48	18.5	125	4	US-08-424-550B-412	Sequence 412, App
39	48	18.5	213	2	US-08-808-550-35	Sequence 35, App1
40	48	18.5	223	4	US-09-465-901-42	Sequence 42, App1
41	48	18.5	307	4	US-09-902-540-15716	Sequence 15716, A
42	48	18.5	315	4	US-09-543-681A-7121	Sequence 7121, App
43	48	18.5	367	4	US-09-236-840A-3	Sequence 3, App11
44	48	18.5	376	4	US-09-721-870-113	Sequence 113, App
45	47.5	18.3	106	4	US-09-252-991A-29961	Sequence 29961, A

ALIGNMENTS

```
RESULT 1
US-09-482-273-180
; Sequence 180, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; EARLIER FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 180
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-482-273-180

Query Match      100.0%; Score 260; DB 4; Length 49;
Best local similarity 100.0%; Pred. No. 6.8e-29;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 MNLIGMIFSMCGMLKLMKCAWVAVYCSFISFANSRSSEDTKOMSSFM 49
Db      1 MNLIGMIFSMCGMLKLMKCAWVAVYCSFISFANSRSSEDTKOMSSFM 49

RESULT 2
US-09-188-930-313
; Sequence 313, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Muriel, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
```

```
/ NUMBER OF SEQ ID NOS: 348
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 313
/ LENGTH: 70
/ TYPE: PRT
/ ORGANISM: Human
US-09-188-930-313
```

```
Query Match
Best Local Similarity 100.0%; Score 260; DB 3; Length 70;
Pred. No. 1e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MNLIGMIFSMCGMLKTKKCAWVAAYCSFISFANSRSSSEDTKOMSSFM 49
DB 1 MNLIGMIFSMCGMLKTKKCAWVAAYCSFISFANSRSSSEDTKOMSSFM 49
```

```
RESULT 3
US-09-312-283C-313
/ Sequence 313, Application US/09312283C
/ Patent No. 6573095
/ GENERAL INFORMATION:
/ APPLICANT: Watson, James D.
/ APPLICANT: Strachan, Lorna
/ APPLICANT: Sleeman, Matthew
/ APPLICANT: Onrust, Rene
/ APPLICANT: Marison, James G.
/ APPLICANT: Kumble, Krishanand D.
/ TITLE OF INVENTION: Compositions Isolated from Skin Cells
/ FILE REFERENCE: 11000.1011c2
/ CURRENT APPLICATION NUMBER: US/09/312,283C
/ CURENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 425
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 313
/ LENGTH: 70
/ TYPE: PRT
/ ORGANISM: Mouse
US-09-312-283C-313
```

```
Query Match
Best Local Similarity 100.0%; Score 260; DB 4; Length 70;
Pred. No. 1e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MNLIGMIFSMCGMLKTKKCAWVAAYCSFISFANSRSSSEDTKOMSSFM 49
DB 1 MNLIGMIFSMCGMLKTKKCAWVAAYCSFISFANSRSSSEDTKOMSSFM 49
```

```
RESULT 4
US-09-482-273-116
/ Sequence 116, Application US/09482273
/ Patent No. 6534631
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 71 Human Secreted Proteins
/ FILE REFERENCE: P2030P1
/ CURRENT APPLICATION NUMBER: US/09/482,273
/ CURENT FILING DATE: 2000-01-13
/ EARLIER APPLICATION NUMBER: PCT/US99/15849
/ EARLIER FILING DATE: 1999-07-14
/ EARLIER APPLICATION NUMBER: 60/092,921
/ EARLIER FILING DATE: 1998-07-15
/ EARLIER APPLICATION NUMBER: 60/092,922
/ EARLIER FILING DATE: 1998-07-15
/ EARLIER APPLICATION NUMBER: 60/092,956
/ EARLIER FILING DATE: 1998-07-15
/ NUMBER OF SEQ ID NOS: 267
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 116
/ LENGTH: 71
/ TYPE: PRT
```

```
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (71)
/ OTHER INFORMATION: Xaa equals stop translation
US-09-482-273-116
```

```
Query Match
Best Local Similarity 100.0%; Score 260; DB 4; Length 71;
Pred. No. 1e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MNLIGMIFSMCGMLKTKKCAWVAAYCSFISFANSRSSSEDTKOMSSFM 49
DB 1 MNLIGMIFSMCGMLKTKKCAWVAAYCSFISFANSRSSSEDTKOMSSFM 49
```

```
RESULT 5
US-09-247-155-90
/ Sequence 90, Application US/09247155A
/ Patent No. 6312822
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, Jean-Baptiste
/ APPLICANT: Duclet, Aymeric
/ APPLICANT: Bouqueloret, Lydie
/ TITLE OF INVENTION: Complementary DNAs
/ FILE REFERENCE: GENSET.021A
/ CURRENT APPLICATION NUMBER: US/09/247,155A
/ CURENT FILING DATE: 1999-02-09
/ EARLIER APPLICATION NUMBER: 60/074,121
/ EARLIER FILING DATE: 1998-02-09
/ EARLIER APPLICATION NUMBER: 60/081,563
/ EARLIER FILING DATE: 1998-04-13
/ EARLIER APPLICATION NUMBER: 60/096,116
/ EARLIER FILING DATE: 1998-08-10
/ EARLIER APPLICATION NUMBER: 60/099,273
/ EARLIER FILING DATE: 1998-10-04
/ NUMBER OF SEQ ID NOS: 182
/ SOFTWARE: Patent.pm
/ SEQ ID NO 90
/ LENGTH: 106
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: -71..-1
US-09-247-155-90
```

```
Query Match
Best Local Similarity 100.0%; Score 260; DB 3; Length 106;
Pred. No. 1.6e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MNLIGMIFSMCGMLKTKKCAWVAAYCSFISFANSRSSSEDTKOMSSFM 49
DB 37 MNLIGMIFSMCGMLKTKKCAWVAAYCSFISFANSRSSSEDTKOMSSFM 85
```

```
RESULT 6
US-09-513-999C-8115
/ Sequence 8115, Application US/09513999C
/ Patent No. 6783961
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Duclet, A.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
/ FILE REFERENCE: 59, US2, REG
/ CURRENT APPLICATION NUMBER: US/09/513,999C
/ CURENT FILING DATE: 2000-02-24
/ PRIOR FILING DATE: 1999-02-26
/ NUMBER OF SEQ ID NOS: 36681
/ SOFTWARE: Patent.pm
```


SEQ ID NO 8115
LENGTH: 106
TYPE: PRT
ORGANISM: Homo sapiens
US-09-513-999C-8115

Query Match 100.0%; Score 260; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.6e-28;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMGMLKXKMCAMVAVYCSFISFANSRSSEDTKQWMSFM 49
DB 37 MNLIGMIFSMGMLKXKMCAMVAVYCSFISFANSRSSEDTKQWMSFM 85

RESULT 7

US-09-270-767-33341
Sequence 33341, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 33341
LENGTH: 114
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-33341

Query Match 83.5%; Score 217; DB 4; Length 114;
Best Local Similarity 77.6%; Pred. No. 1.6e-22;
Matches 38; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMGMLKXKMCAMVAVYCSFISFANSRSSEDTKQWMSFM 49
DB 43 MNLIGMIFSMGMLKXKMCAMFALYCSCISFANSRSRSDAKOVLSSFM 91

RESULT 8

US-09-270-767-48558
Sequence 48558, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 48558
LENGTH: 114
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-48558

Query Match 83.5%; Score 217; DB 4; Length 114;
Best Local Similarity 77.6%; Pred. No. 1.6e-22;
Matches 38; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMGMLKXKMCAMVAVYCSFISFANSRSSEDTKQWMSFM 49
DB 43 MNLIGMIFSMGMLKXKMCAMFALYCSCISFANSRSRSDAKOVLSSFM 91

RESULT 9
US-09-270-767-39637
Sequence 39637, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 39637
LENGTH: 240
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39637

Query Match 23.5%; Score 61; DB 4; Length 240;
Best Local Similarity 24.1%; Pred. No. 1.4;
Matches 13; Conservative 10; Mismatches 21; Indels 10; Gaps 1;

QY 2 NULGMIFSMGMLKXKMCAMVAVYCSFISF-----ANSRSSEDTKQW 45
DB 38 NICTVFAMCGXKTKIKNGCLIVLYCLVFFFXAIWQDKYXNENSDSYEQFL 91

RESULT 10

US-09-270-767-54854
Sequence 54854, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 54854
LENGTH: 240
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54854

Query Match 23.5%; Score 61; DB 4; Length 240;
Best Local Similarity 24.1%; Pred. No. 1.4;
Matches 13; Conservative 10; Mismatches 21; Indels 10; Gaps 1;

QY 2 NULGMIFSMGMLKXKMCAMVAVYCSFISF-----ANSRSSEDTKQW 45
DB 38 NICTVFAMCGXKTKIKNGCLIVLYCLVFFFXAIWQDKYXNENSDSYEQFL 91

RESULT 11

US-09-254-227A-5
Sequence 5, Application US/09254227A
Patent No. 6696257
GENERAL INFORMATION:
APPLICANT: Ahmad, Sultan
APPLICANT: Banville, Denis
APPLICANT: Portin, Yves
APPLICANT: Lembo, Paola
APPLICANT: O'Donnell, Dajan
APPLICANT: Shi-Hsiang, Shen
TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
FILE REFERENCE: 81823/268117
CURRENT APPLICATION NUMBER: US/09/254,227A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 5
LENGTH: 322

TYPE: PRT
ORGANISM: Homo sapiens
US-09-254-227A-5

Query Match 22.5%; Score 58.5; DB 4; Length 322;
Best Local Similarity 25.0%; Pred. No. 4.2;
Matches 14; Conservative 10; Mismatches 19; Indels 13; Gaps 2;

Qy 6 MFSCGMLKMKWCA-----WVAVYCS-----FISFANSSSDTKQMMSSF 48
Db 223 LVFLGCLPFGIQWALFSRIHLDKVLFCHVHLVSIPLSALNSSANPLIYFFVGSF 278

RESULT 12
US-09-254-227A-3
Sequence 3, Application US/09254227A
Patent No. 6696257
GENERAL INFORMATION:
APPLICANT: Ahmad, Sultan
APPLICANT: Banville, Denis
APPLICANT: Fortin, Yves
APPLICANT: Lembo, Paola
APPLICANT: O'Donnell, Dajan
APPLICANT: Shi-Heiang, Shen
TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
FILE REFERENCE: 81823/268117
CURRENT APPLICATION NUMBER: US/09/254,227A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 322
TYPE: PRT
ORGANISM: Homo sapiens
US-09-254-227A-3

Query Match 21.0%; Score 54.5; DB 4; Length 322;
Best Local Similarity 23.2%; Pred. No. 15;
Matches 13; Conservative 11; Mismatches 19; Indels 13; Gaps 2;

Qy 6 MFSCGMLKMKWCA-----WVAVYCS-----FISFANSSSDTKQMMSSF 48
Db 223 LVFLGCLPFGIQWALFSRIHLDKVLFCHVHLVSIPLSALNSSANPLIYFFVGSF 278

RESULT 13
US-09-465-901-16
Sequence 16, Application US/09465901
Patent No. 6492143
GENERAL INFORMATION:
APPLICANT: Reed, Randall
APPLICANT: Yau, King-Wai
APPLICANT: Krautwurst, Dietmar
TITLE OF INVENTION: Olfactory Receptor Expression Libraries
TITLE OF INVENTION: ad Methods of Making and Using Them
FILE REFERENCE: 001107 00105
CURRENT APPLICATION NUMBER: US/09/465,901
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: 60/112,605
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 57
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 223
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
OTHER INFORMATION: PCR primer
US-09-465-901-16

Query Match 20.4%; Score 53; DB 4; Length 223;
Best Local Similarity 31.8%; Pred. No. 16;
Matches 14; Conservative 12; Mismatches 16; Indels 2; Gaps 1;

Qy 8 FSCGMLKMKWCAWVAVYCSFISFANSSSDTK--QMMSSFM 49
Db 175 FSTCSSHLIVSVSYSSVFCAYSPASSYSPEKSKVTSLVYBFL 218

RESULT 14
US-09-248-796A-25087
Sequence 25087, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 25087
LENGTH: 65
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-25087

Query Match 20.0%; Score 52; DB 4; Length 65;
Best Local Similarity 35.3%; Pred. No. 5.5;
Matches 12; Conservative 4; Mismatches 18; Indels 0; Gaps 0;

Qy 13 LMLKMKWCAWVAVYCSFISFANSSSDTKQMS 46
Db 6 LKDLKMGKTKCNVCRFPFKKQKTKTKTKTS 39

RESULT 15
US-09-248-796A-15809
Sequence 15809, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 15809
LENGTH: 190
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-15809

Query Match 20.0%; Score 52; DB 4; Length 190;
Best Local Similarity 28.6%; Pred. No. 18;
Matches 12; Conservative 9; Mismatches 13; Indels 8; Gaps 1;

Qy 5 GMFSCGMLKMKWCAWVAVYCSFISFANSSSDTKQMS 46
Db 132 GILFTCC-----AQFVNLICLFVYISKRSSRDLKSLPS 165

Search completed: March 6, 2005, 08:10:26
Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 08:08:21 ; Search time 93 Seconds

(Without alignments)
173.368 Million cell updates/sec

Title: US-10-664-358-549

Perfect score: 260
Sequence: 1 MNLMGIFMNCGLMKLKWKC.....ISFANSSSDTQMMSSFM 49Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 329044822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum March 04
Maximum March 1004

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubppaa/PTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubppaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubppaa/US10E_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	260	100.0	49 9 US-09-984-245-187	Sequence 187, App
2	260	100.0	49 10 US-09-984-276-180	Sequence 180, App
3	260	100.0	49 10 US-09-984-271-180	Sequence 180, App
4	260	100.0	49 10 US-09-966-262-187	Sequence 187, App
5	260	100.0	49 10 US-09-983-966-187	Sequence 187, App
6	260	100.0	49 14 US-10-059-395-187	Sequence 187, App
7	260	100.0	49 14 US-10-143-090-187	Sequence 187, App
8	260	100.0	70 10 US-09-866-050A-313	Sequence 313, App
9	260	100.0	71 10 US-09-884-276-116	Sequence 116, App
10	260	100.0	71 10 US-09-984-271-116	Sequence 116, App
11	260	100.0	106 10 US-09-903-190-90	Sequence 90, Appl
12	260	100.0	106 11 US-09-978-160A-746	Sequence 746, App
13	260	100.0	106 15 US-10-264-237-2068	Sequence 2068, Ap

14	260	100.0	112 9 US-09-925-301-1413	Sequence 1413, Ap
15	260	100.0	132 15 US-10-276-774-2348	Sequence 2348, Ap
16	158	60.8	30 9 US-09-984-245-324	Sequence 324, App
17	158	60.8	30 10 US-09-966-262-324	Sequence 324, App
18	158	60.8	30 10 US-09-983-966-324	Sequence 324, App
19	158	60.8	30 14 US-10-059-395-324	Sequence 324, App
20	158	60.8	30 14 US-10-143-090-324	Sequence 324, App
21	97	37.3	20 9 US-09-984-245-322	Sequence 322, App
22	97	37.3	20 10 US-09-966-262-322	Sequence 322, App
23	97	37.3	20 10 US-09-983-966-322	Sequence 322, App
24	97	37.3	20 14 US-10-059-395-322	Sequence 322, App
25	97	37.3	20 14 US-10-143-090-322	Sequence 322, App
26	90.5	34.8	126 15 US-10-425-114-64499	Sequence 64499, A
27	90.5	34.8	153 16 US-10-767-701-40770	Sequence 40770, A
28	86.5	33.3	113 16 US-10-437-963-158231	Sequence 158231, A
29	86.5	33.3	118 15 US-10-425-114-67570	Sequence 67570, A
30	86.5	33.3	120 15 US-10-425-114-45342	Sequence 45342, A
31	86.5	33.3	132 15 US-10-425-114-67236	Sequence 67236, A
32	74.5	28.7	101 15 US-10-424-599-265641	Sequence 265641, A
33	74.5	28.7	118 15 US-10-425-114-49227	Sequence 49227, A
34	61	23.5	360 15 US-10-104-047-3609	Sequence 3609, Ap
35	61	23.5	406 15 US-10-220-475A-18	Sequence 47, Appl
36	61	23.5	406 15 US-10-072-012-447	Sequence 447, Appl
37	61	23.5	415 15 US-10-072-012-122	Sequence 122, App
38	61	23.5	416 15 US-10-072-012-124	Sequence 124, App
39	59	22.7	432 15 US-10-369-493-1081	Sequence 1081, Ap
40	58	22.3	391 15 US-10-104-047-2798	Sequence 2798, Ap
41	57	21.9	398 16 US-10-437-963-145446	Sequence 145446, A
42	56.5	21.7	80 16 US-10-437-963-151241	Sequence 151241, A
43	56.5	21.7	539 16 US-10-437-963-154140	Sequence 154140, A
44	56.5	21.7	553 16 US-10-437-963-147568	Sequence 147568, A
45	55.5	21.3	511 14 US-10-080-170-641	Sequence 641, App

ALIGNMENTS

RESULT 1
US-09-984-245-187
Sequence 187, Application US/09984245
Patent No. US20020165374A1
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OR INVENTION: 87 Human Secreted Proteins
FILE REFERENCE: P2004P1
CURRENT FILING DATE: 2001-10-29
CURRENT APPLICATION NUMBER: US/09/984,245
PRIOR APPLICATION NUMBER: 1998-09-17
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: US 60/041,276
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: US 60/042,344
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: US 60/048,094
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/048,350
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/048,188
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/048,135
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/050,937
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/048,187
PRIOR FILING DATE: 1997-05-30
PRIOR APPLICATION NUMBER: US 60/048,099
PRIOR FILING DATE: 1997-05-30

```

; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 187
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-245-187

```

```

Query Match          100.0%; Score 260; DB 9; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSEDTKOMSSFM 49
Db 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSEDTKOMSSFM 49

```

```

RESULT 2
US-09-984-276-180
; Sequence 180, Application US/09984276
; Publication No. US20030017500A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,276
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 180
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-276-180

```

```

Query Match          100.0%; Score 260; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSEDTKOMSSFM 49
Db 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSEDTKOMSSFM 49

```

```

RESULT 3
US-09-984-271-180
; Sequence 180, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 180
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-271-180

```

```

Query Match          100.0%; Score 260; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSEDTKOMSSFM 49
Db 1 MNLGMIFSMCGMLKLMKMCAMVAAYCSPISFANSRSSSEDTKOMSSFM 49

```

```

RESULT 4
US-09-966-262-187
; Sequence 187, Application US/09966262
; Publication No. US20030050461A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/966,262
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937

```

```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 187
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-966-262-187

Query Match      100.0%; Score 260; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MNLLGMIFSMCGMLTKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49
Db      1  MNLLGMIFSMCGMLTKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49

RESULT 5
; US-09-983-966-187
; Sequence 187, Application US/09983966
; Publication No. US20030060619A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/983,966
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
```

```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 187
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-983-966-187

Query Match      100.0%; Score 260; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MNLLGMIFSMCGMLTKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49
Db      1  MNLLGMIFSMCGMLTKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49

RESULT 6
; US-10-059-395-187
; Sequence 187, Application US/10059395
; Publication No. US20030018180A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/10/059,395
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US/09/966,262
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
```

```

; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 187
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-059-395-187

Query Match      100.0%; Score 260; DB 14; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNLIGMIFSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKQMMSSFM 49
Db      1 MNLIGMIFSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKQMMSSFM 49

RESULT 7
US-10-143-090-187
; Sequence 187, Application US/10143090
; Publication No. US20030069406a1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004p1
; CURRENT APPLICATION NUMBER: US/10/143,090
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 09/154,707
```

```

; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 187
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-143-090-187

Query Match      100.0%; Score 260; DB 14; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNLIGMIFSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKQMMSSFM 49
Db      1 MNLIGMIFSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKQMMSSFM 49

RESULT 8
US-09-866-050A-313
; Sequence 313, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Muriison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Human
; US-09-866-050A-313

Query Match      100.0%; Score 260; DB 10; Length 70;
Best Local Similarity 100.0%; Pred. No. 6.5e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNLIGMIFSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKQMMSSFM 49
Db      1 MNLIGMIFSMCGMLKMKCAWAVYCSFISFANSRSSSEDTKQMMSSFM 49

RESULT 9
US-09-984-276-116
; Sequence 116, Application US/09984276
; Publication No. US20030017500A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,276
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 116
; LENGTH: 71
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (71)
; OTHER INFORMATION: Xaa equals stop translation
US-09-984-276-116
```

```
Query Match          100.0%; Score 260; DB 10; Length 71;
Best Local Similarity 100.0%; Pred. No. 6.6e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MNLIGMIFSMGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 49
Db      1 MNLIGMIFSMGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 49
```

```

RESULT 10
US-09-984-271-116
; Sequence 116, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (71)
; OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-116
```

```
Query Match          100.0%; Score 260; DB 10; Length 71;
Best Local Similarity 100.0%; Pred. No. 6.6e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MNLIGMIFSMGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 49
Db      1 MNLIGMIFSMGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 49
```

```

RESULT 11
US-09-903-190-90
; Sequence 90, Application US/09903190
; Publication No. US20030162176A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouguelercet, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/903,190
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US/09/247,155A
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/074,121
```

```

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 90
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: ~71...-1
US-09-903-190-90
```

```
Query Match          100.0%; Score 260; DB 10; Length 106;
Best Local Similarity 100.0%; Pred. No. 1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MNLIGMIFSMGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 49
Db      37 MNLIGMIFSMGMLKIKKWCAMVAVYCSFISFANSRSSEDTKQMMSSFM 85
```

```

RESULT 12
US-09-978-360A-746
; Sequence 746, Application US/09978360A
; Publication No. US20040110939A1
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouguelercet, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.USA.CTP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: ~09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 746
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: ~71...-1
US-09-978-360A-746
```

Query Match 100.0%; Score 260; DB 11; Length 106;
Best Local Similarity 100.0%; Pred. No. 1e-26; Indels 0;
Matches 49; Conservative 0; Mismatches 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKTKWCAWVAVYCSFISFANSRSSSEDTKQWSSFM 49
DB 37 MNLIGMIFSMCGMLKTKWCAWVAVYCSFISFANSRSSSEDTKQWSSFM 85

RESULT 13

US-10-264-237-2068
; Sequence 2068, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Biese et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2068
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-2068

Query Match 100.0%; Score 260; DB 15; Length 106;
Best Local Similarity 100.0%; Pred. No. 1e-26; Indels 0;
Matches 49; Conservative 0; Mismatches 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKTKWCAWVAVYCSFISFANSRSSSEDTKQWSSFM 49
DB 37 MNLIGMIFSMCGMLKTKWCAWVAVYCSFISFANSRSSSEDTKQWSSFM 85

RESULT 14

US-09-925-301-1413
; Sequence 1413, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1413
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1413

Query Match 100.0%; Score 260; DB 9; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.1e-26; Indels 0;
Matches 49; Conservative 0; Mismatches 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKTKWCAWVAVYCSFISFANSRSSSEDTKQWSSFM 49
DB 43 MNLIGMIFSMCGMLKTKWCAWVAVYCSFISFANSRSSSEDTKQWSSFM 91

RESULT 15

US-10-276-774-2348

; Sequence 2348, Application US/10276774
; Publication No. US2004003245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-010
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2348
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2348

Query Match 100.0%; Score 260; DB 15; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.3e-26; Indels 0;
Matches 49; Conservative 0; Mismatches 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLKTKWCAWVAVYCSFISFANSRSSSEDTKQWSSFM 49
DB 63 MNLIGMIFSMCGMLKTKWCAWVAVYCSFISFANSRSSSEDTKQWSSFM 111

Search completed: March 6, 2005, 08:19:20
Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 07:47:05 ; Search time 53 Seconds
(without alignments)
88.955 Million cell updates/sec

Title: US-10-664-358-549

Perfect score: 260

Sequence: 1 MNLLGMIFSMGMLKXKWC.....ISFANSRSBDTKQWSSFM 49

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.79:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	203	78.1	113	2	T16603
2	73.5	28.3	107	2	T45621
3	59	22.7	432	2	F64484
4	58	22.3	971	2	E96794
5	58	22.3	426	2	E87488
6	57	21.9	103	2	S59846
7	57	21.9	391	2	G90266
8	57	21.9	489	2	T26382
9	56	21.5	275	2	S61179
10	55.5	21.3	511	2	C70803
11	55	21.2	659	2	T01520
12	54	20.8	520	2	T12487
13	54	20.8	821	2	T19705
14	53.5	20.6	217	2	AG1277
15	53.5	20.6	495	2	D86442
16	53	20.4	209	2	E82388
17	53	20.4	853	2	S34682
18	52.5	20.2	84	2	S62868
19	52.5	20.2	260	2	F81349
20	52	20.0	214	2	G97008
21	52	20.0	231	2	E90887
22	52	20.0	231	2	D85730
23	52	20.0	231	2	S11429
24	52	20.0	523	2	T07834
25	52	20.0	636	2	A86248
26	52	20.0	758	2	T26595
27	51.5	19.8	520	2	T23276
28	51.5	19.8	520	2	S45753
29	51	19.6	283	2	T21622

30	51	19.6	525	2	E86363	probable thiamin b
31	51	19.6	556	1	OXBY34	DNA endonuclease I
32	51	19.6	854	2	AC2309	hypothetical prote
33	50.5	19.4	154	2	G98112	hypothetical prote
34	50.5	19.4	179	2	S38468	hypothetical prote
35	50.5	19.4	217	2	AG1640	hypothetical prote
36	50.5	19.4	229	2	T03405	probable chitinase
37	50.5	19.4	304	2	T33632	probable cytochrome
38	50.5	19.4	484	2	T34251	hypothetical prote
39	50.5	19.4	572	2	T32636	hypothetical prote
40	50.5	19.4	813	2	UC5785	ATP-dependent RNA
41	50	19.2	174	2	D82629	hypothetical prote
42	50	19.2	325	2	S56788	hypothetical prote
43	50	19.2	370	2	H70423	oxygen-independent
44	50	19.2	405	2	E82626	hypothetical prote
45	50	19.2	668	2	E96777	probable anion exc

ALIGNMENTS

RESULT 1
T16603
hypothetical protein K10B2.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16603
R:Miller, N.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans coemid K10B2.
A:Reference number: Z18545
A:Accession: T16603
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-113 <MIL>
A:Cross-references: UNIPROT:Q09993; EMBL:U28730; NID:g860694; PID:g860698; PIDN:AAA68261
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:K10B2.4
A:introns: 15/2; 53/3; 86/2
C:Superfamily: Caenorhabditis elegans hypothetical protein K10B2.4

Query Match
Best local similarity 78.1%; Score 203; DB 2; Length 113;
Matches 34; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNLLGMIFSMGMLKXKWCAMVAVYCSFISFANSRSBDTKQWSSFM 49
DB 38 MNLLGMIFSMGMLKXKWCMLVCSISFANTRSDAKQIVSSFM 86

RESULT 2
T45621
hypothetical protein F13G24.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T45621
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.-J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23009
A:Accession: T45621
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-107 <BEV>
A:Cross-references: UNIPROT:Q9SD88; EMBL:AL133421
A:Experimental source: cultivar Columbia; BAC clone F13G24
C:Genetics:
A:Map position: 5
A:introns: 53/3; 87/2
A:Note: F13G24.160
C:Superfamily: Caenorhabditis elegans hypothetical protein K10B2.4

Query Match 28.3%; Score 73.5; DB 2; Length 107;

A;Reference number: A99139
A;Accession: G90266
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-391 <KUR>
A;Cross-references: UNIPROT:Q97208; GB:AE006641; NID:G13814326; PIDN:AAK41390.1; GSPDB:C
C;Genetics:
A;Gene: S601138

Query Match 21.9%; Score 57; DB 2; Length 391;
Best Local Similarity 23.9%; Pred. No. 6.3;
Matches 11; Conservative 16; Mismatches 19; Indels 0; Gaps 0;

4 IGMIFMCGMLKMKCAWVAVVCSFISFANSRSSEDTKOWMSSEM 49
Db 229 LAVVFGTLTLMVSLGYSPTISAFVAGVAFANSVKNKIKETISITL 274

RESULT 8
T26382
hypothetical protein Y105CSB.f - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26382
R;McMurray, A.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z20208
A;Accession: T26382
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-489 <RTL>
A;Cross-references: UNIPROT:Q9NAM6; EMBL:AL110479; PIDN:CAM54354.1; CESP:Y105CSB.f
A;Experimental source: clone Y105CSB
C;Genetics:
A;Gene: CESP:Y105CSB.f
A;Intons: 58/3; 79/3; 114/1; 148/2; 233/3; 316/3; 432/1
C;Superfamily: *Caenorhabditis elegans* hypothetical protein Y105CSB.f

Query Match 21.9%; Score 57; DB 2; Length 489;
Best Local Similarity 34.6%; Pred. No. 7.7;
Matches 9; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

6 MIFSMGMLKMKCAWVAVVCSFIS 31
Db 87 LVFDGCDLVPIIDICIMWRIFQSFSS 112

RESULT 9
S61179
probable membrane protein YDR384C - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein D9481.20
C;Species: *Saccharomyces cerevisiae*
C;Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C;Accession: S61179; S69668
R;Ding, H.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of *S. cerevisiae* cosmid 9481.
A;Reference number: S61159
A;Accession: S61179
A;Molecule type: DNA
A;Residues: 1-275 <DIN>
A;Cross-references: UNIPROT:Q12359; EMBL:U28373; NID:G849184; PIDN:AA864820.1; PID:G8492
A;Experimental source: strain S288C (AB972)
R;Dietrich, F.S.
submitted to the EMBL Data Library, July 1995
A;Description: The sequence of *S. cerevisiae* cosmid 9481, 9509, 9926, 9461, and lambda
A;Reference number: S69665
A;Accession: S69668
A;Molecule type: DNA
A;Residues: 1-275 <DIE>
A;Cross-references: EMBL:U2274; NID:G927313; PIDN:AA864826.1; PID:G927317; MIPS:YDR384C
C;Genetics:
A;Cross-references: SGD:S0002792

A;Map position: 4R
C;Superfamily: *Saccharomyces cerevisiae* probable membrane protein FUN34
C;Keywords: transmembrane protein
F;89-105/Domain: transmembrane #status predicted <TM1>
F;124-140/Domain: transmembrane #status predicted <TM2>
F;182-198/Domain: transmembrane #status predicted <TM3>
F;208-224/Domain: transmembrane #status predicted <TM4>

Query Match 21.5%; Score 56; DB 2; Length 275;
Best Local Similarity 37.9%; Pred. No. 6.3;
Matches 11; Conservative 4; Mismatches 6; Indels 8; Gaps 1;

7 IFMFCGMLKMKCAWVAVVCSFISFANS 35
Db 242 ILSSC-----CGWYSLYCSVSPSPNS 262

RESULT 10
C70803
hypothetical protein RV3877 - *Mycobacterium tuberculosis* (strain H37Rv)
C;Species: *Mycobacterium tuberculosis*
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: C70803
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgarbes, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70803
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-511 <COL>
A;Cross-references: UNIPROT:O69741; GB:AL022120; GB:AL123456; NID:G3261558; PIDN:CAA1796
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: RV3877

Query Match 21.3%; Score 55.5; DB 2; Length 511;
Best Local Similarity 42.9%; Pred. No. 13;
Matches 9; Conservative 5; Mismatches 4; Indels 3; Gaps 1;

5 GMIFSMGMLKX---KWCAM 22
Db 401 GLITVCGFRSLYAEKWCAM 421

RESULT 11
T01520
hypothetical protein T10M13.18 - *Arabidopsis thaliana*
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01520; T01431
R;Johnson, A.F.; de la Bastide, M.; Lodhi, M.; Hoffman, J.; Hasegawa, A.; Gnoj, L.; Gott
Martensen, R.; McCombie, W.
submitted to the EMBL Data Library, May 1997
A;Description: The sequence of the *Arabidopsis thaliana* T10M13 BAC.
A;Reference number: Z14346
A;Accession: T01520
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-659 <UOH>
A;Cross-references: UNIPROT:O04257; EMBL:AF001308; NID:G2104523; PID:G3912924
A;Experimental source: cultivar Columbia
R;Kallunki, J.; Gibson, A.
submitted to the EMBL Data Library, August 1998
A;Description: The sequence of *A. thaliana* T2H3.
A;Reference number: Z14324
A;Accession: T01431
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 190-659 <KAL>

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 6, 2005, 06:42:10 ; Search time 123 seconds

(without alignments)
203.999 Million cell updates/sec

Title: US-10-664-358-549

Perfect score: 260
Sequence: 1 MNLGIMFSGMLKLMKMC.....ISFANSRSSSDTKQMSSEFM 49Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	100.0	104	2 Q9BV13	Q9BV13 homo sapien
2	260	100.0	106	1 CGEO_HUMAN	Q9Y234 homo sapien
3	260	100.0	106	2 Q6Q7K0	Q6Q7K0 sus scrofa
4	260	100.0	106	2 Q6ZWX0	Q6ZWX0 mus musculu
5	259	99.6	106	2 Q6DKM8	Q6DKM8 xenopus lae
6	259	99.6	106	2 Q6NTW8	Q6NTW8 xenopus lae
7	259	99.6	106	2 Q6PC46	Q6PC46 brachydanio
8	221	85.0	130	2 Q7QIP7	Q7QIP7 anopheles g
9	219	84.2	108	1 U139_MANSE	Q9U516 manduca sex
10	217	83.5	102	2 Q6XIW7	Q6XIW7 drosophila
11	217	83.5	108	1 U139_DROME	Q9V178 drosophila
12	203	78.1	113	1 U139_CABEL	Q09993 caenorhabdi
13	153	34.8	153	2 Q8LKT8	Q8LKT8 sorghum bic
14	86.5	33.3	93	2 Q9AUI6	Q9AUI6 cryza sativ
15	80	30.8	98	2 Q86H65	Q86H65 dictyosceli
16	73.5	28.3	107	1 U139_ARATH	Q9S488 arabidopsis
17	65	25.0	95	1 YA3B_SCHPO	Q09714 schistosach
18	61	23.5	352	2 Q6ICM2	Q6ICM2 homo sapien
19	61	23.5	406	1 S144_HUMAN	Q9UD33 homo sapien
20	61	23.5	415	2 Q6G065	Q6G065 polyangium
21	60	23.1	671	2 Q92KX3	Q92KX3 thizobium m
22	59.5	22.9	439	2 Q6FC66	Q6FC66 ixodes uria
23	59.5	22.9	1802	2 Q9VQ08	Q9VQ08 drosophila
24	59	22.7	209	2 Q639A5	Q639A5 bacillus ce
25	59	22.7	432	1 YE79_MERJA	Q58674 methanococ
26	59	22.7	971	2 Q9C9T5	Q9C9T5 arabidopsic
27	58.5	22.5	322	1 SNS2_HUMAN	Q8D600 homo sapien
28	58	22.3	214	2 Q9N9T4	Q9N9T4 leishmania
29	58	22.3	426	2 Q9A6Z6	Q9A6Z6 caulobacter
30	57.5	22.1	387	2 Q83DP9	Q83DP9 coxiella bu
31	57	21.9	298	2 Q6G1Q0	Q6G1Q0 bartonella

32	57	21.9	358	2 Q8G5K8	Q8G5K8 bifidobacte
33	57	21.9	391	2 Q97Z08	Q97Z08 sulfolobus
34	57	21.9	447	2 Q9NAM8	Q9NAM8 caenorhabdi
35	57	21.9	491	2 Q9U622	Q9U622 drosophila
36	57	21.9	572	2 Q81L79	Q81L79 plasmodium
37	56.5	21.7	503	2 Q7ZYD3	Q7ZYD3 xenopus lae
38	56.5	21.7	543	2 Q7UHR4	Q7UHR4 rhodospirell
39	56.5	21.7	583	2 Q67M66	Q67M66 oryza sativ
40	56.5	21.7	846	2 Q9ZOU6	Q9ZOU6 brassica ol
41	56.5	21.7	1520	2 Q6ZPP4	Q6ZPP4 mus musculu
42	56.5	21.7	4903	1 MLL3_MOUSE	Q8BTH4 mus musculu
43	56	21.5	275	1 YD84_YEAST	Q12359 baccharomyc
44	56	21.5	403	1 S144_MOUSE	Q8R0E9 mus musculu
45	56	21.5	454	2 Q9MJ58	Q9MJ58 loligo blec

ALIGNMENTS

RESULT 1					
ID	Q9BV13	PRELIMINARY;	PRT;	104 AA.	
AC	Q9BV13				
DT	01-JUN-2001 (TREMBLrel. 17, Created)				
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)				
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)				
DE	PBD008 protein (Fragment).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISUE=Eye;				
RA	Strausberg R.;				
RL	Submitted (DBS-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC001192; AA01192.1; -				
DR	InterPro: IPR005351; UPF0139.				
DR	Pfam; PF03669; UPF0139; 1.				
DR	ProDom; PD073147; UPF0139; 1.				
FT	NON TER				
SQ	SEQUENCE	104 AA;	11850 MW;	DOB525EA62800A1 CRC64;	
Query Match					
Best Local Similarity 100.0%; Score 260; DB 2; Length 104;					
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 MNLGIMFSGMLKLMKCAWVAVYCSFISFANSRSSSDTKQMSSEFM 49				
DB	35 MNLGIMFSGMLKLMKCAWVAVYCSFISFANSRSSSDTKQMSSEFM 83				
RESULT 2					
CGEO_HUMAN	STANDARD;	PRT;	106 AA.		
ID	Q9Y256;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	29-MAR-2004 (Rel. 43, Last annotation update)				
DE	UPF0139 protein CGI-140 (PTD008) (My006 protein).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20272150; PubMed=10810093; DOI=10.1101/gr.10.5.703;				
RA	Lai C.-H., Chou C.-Y., Chiang L.-Y., Liu C.-S., Lin W.-C.;				
RT	"Identification of novel human genes evolutionarily conserved in				
RT	Cenohaditis elegans by comparative proteomics."				
RL	Genome Res. 10:703-713 (2000).				
RN	[2]				
RP	SEQUENCE FROM N.A.				

```
RC TISSUE=Pituitary tumor;
RA Song H., Peng Y., Dai M., Huang Q., Mao Y., Zhang Q., Mao M., Fu G.,
RA Luo M., Chen J., Hu R.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA TISSUE=Fetal brain;
RA Mao Y.M., Xie Y., Zhou Z.X.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the UPF0139 (CGI-140) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF151898; AAD34135.1; -
DR EMBL; AF078861; AAD44493.1; -
DR EMBL; AF059620; AAG43119.1; -
DR InterPro: IPR005351; UPF0139.
DR Pfam: PF03669; UPF0139; 1.
DR ProDom: PD073147; UPF0139; 1.
DR Transmembrane.
KW TRANSMEM
FT
SQ SEQUENCE 106 AA; 12068 MW; 7D2BDADF279071B CRC64;
Potential.

Query Match
Best Local Similarity 100.0%; Score 260; DB 1; Length 106;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLIGMIFSMCGIMLKLKMCAMVAAYVCSFISFANSRSSSEDTKQWSSFM 49
37 MNLIGMIFSMCGIMLKLKMCAMVAAYVCSFISFANSRSSSEDTKQWSSFM 85

RESULT 3
ID 0607K0 PRELIMINARY; PRT; 106 AA.
AC 0607K0;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE CGI-140.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN
RP SEQUENCE FROM N.A.
RA Lee H.Y., Cui X.S., Shin M.L., Jeong Y.J., Hwang K.C., Kim N.H.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY553921; AAS76542.1; -
DR InterPro: IPR005351; UPF0139.
DR Pfam: PF03669; UPF0139; 1.
DR ProDom: PD073147; UPF0139; 1.
DR Transmembrane.
KW TRANSMEM
FT
SQ SEQUENCE 106 AA; 11942 MW; 0CB2D4249AD258E7 CRC64;

Query Match
Best Local Similarity 100.0%; Score 260; DB 2; Length 106;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLIGMIFSMCGIMLKLKMCAMVAAYVCSFISFANSRSSSEDTKQWSSFM 49
37 MNLIGMIFSMCGIMLKLKMCAMVAAYVCSFISFANSRSSSEDTKQWSSFM 85

RESULT 4
ID 06ZWX0 PRELIMINARY; PRT; 106 AA.
AC 06ZWX0;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Mus musculus 13 days embryo liver cDNA. RIKEN full-length enriched
DE library, clone:2510002B07 product:PROTEIN CGI-140 (PROTEIN PTD008)
DE homolog (cDNA sequence BC056474).
GN Name=BC056474;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=liver;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RX Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=liver;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RX RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=liver;
RC MEDLINE=11042159; DOI=10.1101/gr.145100;
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugihara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=liver;
RC MEDLINE=11076861; DOI=10.1101/gr.152600;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasaiyagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=liver;
RC Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Aizawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imocani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numata R., Ono M.,
RA Okazaki Y., Okido T., Owa C., Saito R., Saito R., Sakai K., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami K., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
```

RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Tothiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smallus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUG-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK010893; BAB27250.1; -
 DR EMBL; BC056474; AAH56474.1; -
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139.1.
 DR ProDom; PD073147; UPF0139.1.
 SQ SEQUENCE 106 AA; 1206 MM; 7D2BDDAF9279071B CRC64;

Query Match 100.0%; Score 260; DB 2; Length 106;
 Best Local Similarity 100.0%; Pred. No. 8.6e-27;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGMIFFSMCGMLTKKMCANVAVYCSFISFANSRSSEDTKQMSSEFM 49
 DB 37 MNLGMIFFSMCGMLTKKMCANVAVYCSFISFANSRSSEDTKQMSSEFM 85
 ID Q6DKM8 PRELIMINARY; PRT; 106 AA.
 AC Q6DKM8;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE MGC82186 protein.
 GN Name=MGC82186;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Tothiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smallus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT Initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (May-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC071133; AAH71133.1; -
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139.1.
 DR ProDom; PD073147; UPF0139.1.
 SQ SEQUENCE 106 AA; 1185 MM; 169F889AD4BBB90 CRC64;

Query Match 99.6%; Score 259; DB 2; Length 106;
 Best Local Similarity 98.0%; Pred. No. 1.2e-26;
 Matches 48; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGMIFFSMCGMLTKKMCANVAVYCSFISFANSRSSEDTKQMSSEFM 49
 DB 37 MNLGMIFFSMCGMLTKKMCANVAVYCSFISFANSRSSEDTKQMSSEFM 85
 ID Q6NTW8 PRELIMINARY; PRT; 106 AA.
 AC Q6NTW8;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE MGC81480 protein.
 GN Name=MGC81480;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Tothiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smallus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human

RA Nardl J.B.:
RT "Diversity of odorant binding proteins revealed by an expressed
RT sequence tag project on male *Manduca sexta* moth antennae";
RL Insect Mol. Biol. 8:501-518 (1999).
CC -1- SIMILARITY: Belongs to the UPF0139 (CGI-140) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.emb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, AF117572, AAF1664.1; -
CC InterPro: IPR005351; UPF0139.
CC Pfam: PF03669; UPF0139; 1.
CC ProDom: PD073147; UPF0139; 1.
CC K0 Hypothetical protein; Transmembrane.
CC TRANSMEM 77 Potential.
CC SEQUENCE 108 AA; 12125 MW; 1C57B0933A798921 CRC64;
SO
Query Match 84.2%; Score 219; DB 1; Length 108;
Best Local Similarity 77.6%; Pred. No. 2.6e-21;
Matches 38; Conservative 8; Mismatches 3; Indels 0; Gaps 0;
QY 1 MNLGMIFSWCGMLTKKMCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49
DB 34 MNLGMIFSWCGMLTKKMCAMVAVYCSFISFANSRSVDDTKQIVSSFM 82
RESULT 10
Q6XIW7 PRELIMINARY; PRT; 102 AA.
ID Q6XIW7
AC Q6XIW7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to *Drosophila melanogaster* CG10674 (Fragment).
OS *Drosophila yakuba* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
NCBI_TaxID=7245;
OK
RN
RP
RX MEDLINE=22887302; PubMed=14525923; DOI=10.1101/gr.1311003;
RA Domazet-Loso T., Tautz D.;
RT "An evolutionary analysis of orphan genes in *Drosophila*.";
RL Genome Res. 13:2213-2219(2003).
DR EMBL; AY31712; AAR09735.1; -
DR InterPro; IPR005351; UPF0139.
DR Pfam; PF03669; UPF0139; 1.
DR ProDom; PD073147; UPF0139; 1.
FT NON_TER 1
FT NON_TER 1
SO SEQUENCE 102 AA; 11397 MW; B13E02BD0915D79 CRC64;
Query Match 83.5%; Score 217; DB 2; Length 102;
Best Local Similarity 77.6%; Pred. No. 4.6e-21;
Matches 38; Conservative 8; Mismatches 3; Indels 0; Gaps 0;
QY 1 MNLGMIFSWCGMLTKKMCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49
DB 37 MNLGMIFSWCGMLTKKMCAMVAVYCSFISFANSRSASDAKVLSSFM 85
RESULT 11
U139_DROME STANDARD; PRT; 108 AA.
AC Q9VRU8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Hypothetical UPF0139 protein CG10674.
GN ORFNames=CG10674;
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
NCBI_TaxID=7227;
OK
RN
RP
RX STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne U.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hopkins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.F., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Buttle K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Hostin D., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kechum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kul D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.A., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN
RN
RP
RX GENOME REANNOTATION;
RX MEDLINE=22426069; PubMed=12537572;
RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitefield E.J., Bayraktarglu L., Berman B.F.,
RA Beccencourt B.R., Celinker S.E., de Grey A.D.N., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -1- SIMILARITY: Belongs to the UPF0139 (CGI-140) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.emb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----

DR EMBL; AE003567; AAF50797.1; -
 DR InLoc; Q9VRJ8; -
 DR FlyBase; FBgn0035592; CG10674.
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139; 1.
 DR ProDom; PD073147; UPF0139; 1.
 DR Hypothetical protein; Transmembrane.
 FT TRANSMEM 80
 FT Potential.
 SQ SEQUENCE 108 AA; 12082 MW; 440857655D5238DC CRC64;

Query Match 83.5%; Score 217; DB 1; Length 108;
 Best Local Similarity 77.6%; Pred. No. 4.8e-21;
 Matches 38; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTQKMSSFM 49
 DB 37 MNLIGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTQKMSSFM 85

RESULT 12
 U139_CABEL STANDARD; PRT; 113 AA.

AC 009993;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Hypothetical UPF0139 protein K10B2.4 in chromosome II.
 GN ORFNames=K10B2.4;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Polidoridae; Caenorhabditis.
 OX NCBI_Taxid=6239;

RP SEQUENCE FROM N.A.
 RX STRAIN=Bristol N2;
 RX MEDLINE=9069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for investigating biology."
 RL Science 282:2012-2018 (1998).
 CC -1- SIMILARITY: Belongs to the UPF0139 (CGI-140) family.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U28730; AAA68261.1; -
 DR PIR; T16603; T16603
 DR WormBase; WBGene00019607; K10B2.4.
 DR WormPep; K10B2.4; CR02011.
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139; 1.
 DR ProDom; PD073147; UPF0139; 1.
 DR Hypothetical protein; Transmembrane.
 FT TRANSMEM 81
 FT Potential.
 SQ SEQUENCE 113 AA; 12764 MW; 0FEED534A710DD50 CRC64;

Query Match 78.1%; Score 203; DB 1; Length 113;
 Best Local Similarity 69.4%; Pred. No. 3.7e-19;
 Matches 34; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTQKMSSFM 49
 DB 38 MNLIGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTQKMSSFM 86

RESULT 13
 OBLKT8 PRELIMINARY; PRT; 153 AA.

AC Q8LKT8;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Hypothetical protein I70F8.12.
 GN Name=I70F8.12;
 OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoidae; Andropogoneae; Sorghum.
 OX NCBI_Taxid=4558;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=22341483; PubMed=12454082;
 RX Ramakrishna W., Dubcovsky J., Park Y.J., Busso C., Emberton J.,
 RA SamMiguel P., Bennezy J.L.;
 RT "Different types and rates of genome evolution detected by comparative
 RT sequence analysis of orthologous segments from four cereal genomes."
 RL Genetics 162:1389-1400(2002).
 DR EMBL; AF503433; AAM47587.1; -
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139; 1.
 DR ProDom; PD073147; UPF0139; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 153 AA; 16498 MW; 0A59856FB6FC4CE2 CRC64;

Query Match 34.8%; Score 90.5; DB 2; Length 153;
 Best Local Similarity 38.3%; Pred. No. 0.00053;
 Matches 18; Conservative 12; Mismatches 16; Indels 1; Gaps 1;

QY 4 LGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTQKMSSFM 49
 DB 45 LAVFGVGVMLTKYKCSWIAITFCAGSLANKNFENDIKOLSMAFM 91

RESULT 14

Q9AUV6 PRELIMINARY; PRT; 93 AA.

AC Q9AUV6;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-OCT-2003 (TRENBLrel. 17, Last sequence update)
 DE Hypothetical protein OSJNB40040E01.6.
 GN Name=OSJNB40040E01.6;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_Taxid=39947;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Buell C.R., Yuan O., Ouyang S., Moffat K.S., Hill J.N., Burr P.C.,
 RA Heio J., Zismann V., Bai G., Bowman C.L., Fujii C.Y., Vanaken S.E.,
 RA Bowman C.L., Craven B., Uteckack T.R., Khalak H., Feldblum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RA Buell R.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC079887; AAK16184.1; -
 DR Gramene; Q9AUV6; -
 DR InterPro; IPR005351; UPF0139.
 DR Pfam; PF03669; UPF0139; 1.
 DR ProDom; PD073147; UPF0139; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 93 AA; 9961 MW; B92A623A0380C7EC CRC64;

Query Match 33.3%; Score 86.5; DB 2; Length 93;
 Best Local Similarity 36.2%; Pred. No. 0.0011;
 Matches 17; Conservative 12; Mismatches 17; Indels 1; Gaps 1;

QY 4 LGMIFSMCGMLTKMKCAWAVVCSFISFANSSSEDTQKMSSFM 49

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 04:07:08 ; Search time 4556 seconds
(without alignments)
11039.621 Million cell updates/sec

Title: US-10-664-358-36

Perfect score: 1038

Sequence: 1 gnaatcgagcagcagcgttaac.....aaaaaaaaaaaaaacgcga 1038

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ha:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_gy:*
13: gb_un:*
14: gb_vt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1035	99.7	1038	6 BD074351	BD074351 50 human
2	1011.6	97.5	125780	2 AL138904	AL138904 Homo sapi
3	1011.6	97.5	141670	2 AC104170	AC104170 Homo sapi
4	1011.6	97.5	164500	2 AL355968	AL355968 Homo sapi
5	1011.6	97.5	169304	2 AL731898	AL731898 Homo sapi
6	207.8	20.0	41322	2 AC004076	AC004076 Homo sapi
7	206.2	19.9	251124	9 HUAE000660	AE000660 Homo sapi
8	205.4	19.8	99108	9 AC008666	AC008666 Homo sapi
9	205.4	19.8	176629	9 AC008641	AC008641 Homo sapi
10	203.8	19.6	205819	2 AC106805	AC106805 Homo sapi
11	203.6	19.6	116368	9 HS112F19	AL034420 Human DNA
12	203	19.6	159490	9 AC110594	AC110594 Homo sapi
13	202.8	19.5	123902	9 AC108026	AC108026 Homo sapi
14	202.4	19.5	133860	9 AL157709	AL157709 Human DNA
15	202.2	19.5	65335	9 AC112698	AC112698 Homo sapi
16	202.2	19.5	162104	9 AC026325	AC026325 Homo sapi
17	202.2	19.5	176599	2 AC114911	AC114911 Homo sapi
18	202	19.5	181188	2 AC067983	AC067983 Homo sapi
19	202	19.5	194385	9 CNS00001	AL049776 Human chr

20	201.6	19.4	172770	2 AC023133	AC023133 Homo sapi
21	201.6	19.4	173053	2 AC129893	AC129893 Homo sapi
22	201.6	19.4	181517	2 AC126327	AC126327 Homo sapi
23	201.4	19.4	29133	2 AC139398	AC139398 Homo sapi
24	201.4	19.4	50068	2 AC139097	AC139097 Homo sapi
25	201.4	19.4	134059	2 AC016284	AC016284 Homo sapi
26	201.4	19.4	165902	2 AC106763	AC106763 Homo sapi
27	201.4	19.4	176104	9 AC007163	AC007163 Homo sapi
28	201.4	19.4	206302	2 AC131125	AC131125 Homo sapi
29	201.2	19.4	76771	9 AL451051	AL451051 Human DNA
30	200.6	19.3	168839	9 AC022826	AC022826 Homo sapi
31	200.4	19.3	74486	9 AC055764	AC055764 Homo sapi
32	200.4	19.3	96202	9 AL138824	AL138824 Human DNA
33	200.4	19.3	150140	9 AC112216	AC112216 Homo sapi
34	200.2	19.3	160378	9 AC010281	AC010281 Homo sapi
35	200	19.3	68196	6 C0870013	C0870013 Sequence
36	200	19.3	138121	9 AC025431	AC025431 Homo sapi
37	200	19.3	209779	9 AC066616	AC066616 Homo sapi
38	199.8	19.2	121158	9 AC092977	AC092977 Homo sapi
39	199.8	19.2	122748	9 HS742C19	AL031846 Human DNA
40	199.8	19.2	181636	2 AC025108	AC025108 Homo sapi
41	199.6	19.2	35001	9 AC139143	AC139143 Homo sapi
42	199.6	19.2	152156	9 AC018552	AC018552 Homo sapi
43	199.6	19.2	156503	9 AC011511	AC011511 Homo sapi
44	199.6	19.2	161345	2 AC091968	AC091968 Homo sapi
45	199.6	19.2	169786	9 BS000112	BS000112 Pan trogl

ALIGNMENTS

RESULT 1	BD074351	1038 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD074351	50 human secreted proteins.			
DEFINITION	BD074351	50 human secreted proteins.			
ACCESSION	BD074351.1	GI:22619954			
VERSION	JP 2001514024-A/56.				
KEYWORDS	JP 2001514024-A/56.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryote, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Lagomorpha; Carnivora; Homnidae; Homo. 1 (bases 1 to 1038)				
AUTHORS	Moore, P.A., Ruben, S.M., Lafleur, D.W., Shi, Y., Rosen, C.A., Olsen, H.S., Ebner, R. and Brewer, L.A.				
TITLE	50 human secreted proteins				
JOURNAL	Patent: JP 2001514024-A 56 11-SEP-2001; HUMAN GENOME SCIENCES INC				
COMMENT	OS Homo sapiens (human) PN JP 2001514024-A/56 PD 11-SEP-2001 PF 03-SEP-1998 JP 2000508394 PR 05-SEP-1997 US 60/057669, 05-SEP-1997 US 60/057663 PR 05-SEP-1997 US 60/057669, 12-SEP-1997 US 60/058667 PR 12-SEP-1997 US 60/058974, 12-SEP-1997 US 60/058973 PR 12-SEP-1997 US 60/058666 PI PAUL A MOORE, STEVEN W RUBEN, DAVID W LAFLEUR, YANGGU SHI, CRAIG A				
	PI ROSEN, PI HENRIK S OLSEN, REINHARD EBNER, LAURIE A BREWER PC C12N15/09, A61K38/00, A61K48/00, A61P17/02, A61P19/06, A61P19/00, PC A61P19/02, PC A61F29/00, A61P31/04, A61P31/12, A61P35/00, C07K14/435, C07K16/18, PC C12N1/15, PC C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, G01N33/53, C12N15/ PC 00, PC A61K37/02, C12N5/00 CC n equals a,t,g, or c FH key FT source FT 1. 1038 FT /organism='Homo sapiens (human)'. FEATURES Location/Qualifiers				

source 1..1038

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.7%; Score 1035; DB 6; Length 1038;
Best Local Similarity 100.0%; Pred. No. 1.4e-191;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GNAATTCGCGACGAGTTAATGTAATAAATATTTCTAATAATGAAATTTAAATGGAATTGA 60
DB 1 GNAATTCGCGACGAGTTAATGTAATAAATATTTCTAATAATGAAATTTAAATGGAATTGA 60
QY 61 GCATCATGAAAAAAGCTCTTACCTGTGAAAAACATATTGTGTACATTTGGTCACTA 120
DB 61 GCATCATGAAAAAAGCTCTTACCTGTGAAAAACATATTGTGTGTCAATTTGGTCAACTA 120
QY 121 ATCTTCAATTAATCTTTAGTAACTAATAGTTAAGTTAGACAGAGTGGCAGTCTTATATAG 180
DB 121 ATCTTCAATTAATCTTTAGTAACTAATAGTTAAGTTAGACAGTGGCAGTCTTATATAG 180
QY 181 TAAATGCGAGCTGACGACATGAATAATACATATCTAATATTTTGTGACTATCTTATAG 240
DB 181 TAAATGCGAGCTGACGACATGAATAATACATATCTAATATTTTGTGACTATCTTATAG 240
QY 241 AAAATCAGAGAAATTCAAAACCTGTGAGTTTGGGATAGTGCACATTTTAAATAGT 300
DB 241 AAAATCAGAGAAATTCAAAACCTGTGAGTTTGGGATAGTGCACATTTTAAATAGT 300
QY 301 GCGGATATTTTATACATGATTTGAGCTTTGTGMAAATATTTTCCCTGACCTTTTATTTA 360
DB 301 GCGGATATTTTATACATGATTTGAGCTTTGTGMAAATATTTTCCCTGACCTTTTATTTA 360
QY 361 GATGAGATTAACAGTGTAGGCAACTTATATATCTGCAACTCCATTAATGTCATAGTC 420
DB 361 GATGAGATTAACAGTGTAGGCAACTTATATATCTGCAACTCCATTAATGTCATAGTC 420
QY 421 AGACATATCCCATCTAAATATATATGTTGTAAATAGCTTTGTAATAGTGTG 480
DB 421 AGACATATCCCATCTAAATATATATGTTGTAAATAGCTTTGTAATAGTGTG 480
QY 481 AGGTCATTATATCAACAAGCTCTTCAAGGKATTCATTATATAAAACCTTGKTTTATCTTG 540
DB 481 AGGTCATTATATCAACAAGCTCTTCAAGGKATTCATTATATAAAACCTTGKTTTATCTTG 540
QY 541 TGAATAMCGTTTTTTCATGCAAGTTAAATTTCTGAGCCTTTAATTTTATTAAT 600
DB 541 TGAATAMCGTTTTTTCATGCAAGTTAAATTTCTGAGCCTTTAATTTTATTAAT 600
QY 601 ATATAAGATGATGATGATATGATCAACAATAACAGAAAAATTAACAGATTTGCTGTG 660
DB 601 ATATAAGATGATGATGATATGATCAACAATAACAGAAAAATTAACAGATTTGCTGTG 660
QY 661 GCTTTTGTCTAATATGTTACTGTGACAAATCTTAGCCAGTCTTCAATTTGCTTTGAGAT 720
DB 661 GCTTTTGTCTAATATGTTACTGTGACAAATCTTAGCCAGTCTTCAATTTGCTTTGAGAT 720
QY 721 GAAAGTACTTGAATTTAGTTCAGAGGGGCTGGGCGGATAGCTGATCCCTGTGCTCCAGTG 780
DB 721 GAAAGTACTTGAATTTAGTTCAGAGGGGCTGGGCGGATAGCTGATCCCTGTGCTCCAGTG 780
QY 781 CTTTGCGGGGCGAGGCGAGTGTGATCACTTAAGGTCAAGATTTGAGACCAAGCTGCGCA 840
DB 781 CTTTGCGGGGCGAGGCGAGTGTGATCACTTAAGGTCAAGATTTGAGACCAAGCTGCGCA 840
QY 841 ACATGCTGAAACGTTGTCTTCTAATAAATCAAAAATTAGACAGGCGTGTGTCACACA 900
DB 841 ACATGCTGAAACGTTGTCTTCTAATAAATCAAAAATTAGACAGGCGTGTGTCACACA 900
QY 901 TCTGTAAATTCAGGCTACTCAGAGGCTTAACAAGAAAAATTCCTGAACCTGGGAGGAGCAG 960
DB 901 TCTGTAAATTCAGGCTACTCAGAGGCTTAACAAGAAAAATTCCTGAACCTGGGAGGAGCAG 960
```

```
QY 961 AGGTTCAGTGAAGCCATTGACCTCCAGCCTGGGGCAACAGTAGACTTTGTCTCAAAA 1020
DB 961 AGGTTCAGTGAAGCCATTGACCTCCAGCCTGGGGCAACAGTAGACTTTGTCTCAAAA 1020
QY 1021 AAAAAAAAAAAACTCCGA 1038
DB 1021 AAAAAAAAAAAACTCCGA 1038
```

RESULT 2

AL138904/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Plumb, B.

Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Aug 27, 2000 this sequence version replaced gi:9796269.

----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquerry@sanger.ac.uk

----- Project Information

Center project name: dj932F4

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 118782 bases at least Q40

Consensus quality: 121247 bases at least Q30

Consensus quality: 122835 bases at least Q20

Insert size: 125380; sum-of-contigs

Insert size: 139370; 4.2% error; agarose-fp

Quality coverage: 3.66x in Q20 bases; sum-of-contigs Quality

coverage: 3.42x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently

* consists of 15 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

```
* 1 3279: contig of 3279 bp in length
* 3280 3379: gap of 100 bp
* 3380 13996: contig of 10617 bp in length
* 13997 14096: gap of 100 bp
* 14097 17758: contig of 3662 bp in length
* 17759 17858: gap of 100 bp
* 17859 22625: contig of 4767 bp in length
* 22626 22725: gap of 100 bp
* 22726 33140: contig of 10415 bp in length
* 33141 33240: gap of 100 bp
* 33241 36089: contig of 2849 bp in length
* 36090 36189: gap of 100 bp
* 36190 43717: contig of 7527 bp in length
* 43717 43817: gap of 100 bp
* 43817 50092: contig of 6276 bp in length
* 50093 50192: gap of 100 bp
* 50193 56140: contig of 5948 bp in length
```

* 56141 56240: gap of 100 bp
* 56241 59540: contig of 3300 bp in length
* 59541 59640: gap of 100 bp
* 59641 82808: contig of 23168 bp in length
* 82809 82908: gap of 100 bp
* 82909 100070: contig of 17162 bp in length
* 100071 100170: gap of 100 bp
* 100171 116654: contig of 16484 bp in length
* 116655 116754: gap of 100 bp
* 116755 119341: contig of 2587 bp in length
* 119342 119442: gap of 100 bp
* 119442 126780: contig of 7339 bp in length.
Location/Qualifiers
1. 126780
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="P32.3-33"
/clone="RP5-932F4"
/clone_1fb="RPCT-5"
1. 3279
/note="assembly fragment:00163
fragment chain:1"
3380. 13996
/note="assembly fragment:00030
fragment chain:1"
14097. 17758
/note="assembly fragment:00986
fragment chain:1"
17859. 22625
/note="assembly fragment:01111
fragment chain:1"
22726. 33140
/note="assembly fragment:00681
fragment chain:2"
33241. 36089
/note="assembly fragment:01055
fragment chain:2"
36190. 43716
/note="assembly fragment:00252
fragment chain:2"
43817. 50092
/note="assembly fragment:00161
fragment chain:2"
50193. 56140
/note="assembly fragment:00987
fragment chain:3"
56241. 59540
/note="assembly fragment:00088
fragment chain:3"
59641. 82808
/note="assembly fragment:00044"
82909. 100070
/note="assembly fragment:00064"
100171. 116654
/note="assembly fragment:00287"
116755. 119341
/note="assembly fragment:00337.0"
119442. 126780
/note="assembly fragment:00940"

FEATURES
source
Query Match 97.5%; Score 1011.6; DB 2; Length 126780;
Best Local Similarity 99.1%; Pred. No. 3.5e-187;
Matches 1011; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
ORIGIN
Query 16 TTAATGTAATAAATTTCTAATGAATTTAATGGAATTAGACATTCATAGAAAAA 75
Db 64760 TTAATGTAATAAATTTCTAATGAATTTAATGGAATTAGACATTCATAGAAAAA 64701
Gy 76 TGCTCTACTGTGTAACATTAATTTGTACATTTTGTCACATGATCTTCAATTACTT 135
|||||

Db 64700 TGCTCTACTGTGTAACATTAATTTGTACATTTTGTCACATGATCTTCAATTACTT 64641
Gy 136 TTAATGTAATAAATTTCTAATGAATTTAATGGAATTAGACATTCATAGAAAAA 195
Db 64640 TTAATGTAATAAATTTCTAATGAATTTAATGGAATTAGACATTCATAGAAAAA 64581
Gy 196 AGCATGAAAAATTAACATATCTAATATTTTGTGACATCTATCTTAATAGAAAAATCAGAAATTT 255
Db 64580 AGCATGAAAAATTAACATATCTAATATTTTGTGACATCTATCTTAATAGAAAAATCAGAAATTT 64521
Gy 256 CAAACCTGTGTAGTTTGTAGGGTATAGTCATATTAATAATGTGCGGTATATTTATAC 315
Db 64520 CAAACCTGTGTAGTTTGTAGGGTATAGTCATATTAATAATGTGCGGTATATTTATAC 64461
Gy 316 ATGATTTAGAGTTTGTGMAAATATTTCCCGGACCTTTAATTTAGATGAATCAGT 375
Db 64460 ATGATTTAGAGTTTGTGMAAATATTTCCCGGACCTTTAATTTAGATGAATCAGT 64401
Gy 376 GTAGCAAACTTATATATATCTGTCACTCCATTAAGTGCATAGTCAGACTCCCATG 435
Db 64400 GTAGCAAACTTATATATATCTGTCACTCCATTAAGTGCATAGTCAGACTCCCATG 64341
Gy 436 CTAATAATTATAGTTGTAAAAATACGCTTTGTAAATAGTTGTAGGTCATTAACCA 495
Db 64340 CTAATAATTATAGTTGTAAAAATACGCTTTGTAAATAGTTGTAGGTCATTAACCA 64281
Gy 496 AGCTTCAGAGKATTAACATTAATAAACCTTGCTTTATCTGTGAATMCGGTTTT 555
Db 64280 AGCTTCAGAGTATTAACATTAATAAACCTTGCTTTATCTGTGAATMCGGTTTT 64221
Gy 556 TCCATGCMAAGTTAAAAATTTCTCAGCCTTTAATTTTATTAATATATTAAGATGAT 615
Db 64220 TCCATGCMAAGTTAAAAATTTCTCAGCCTTTAATTTTATTAATATATTAAGATGAT 64161
Gy 616 GAGTATGACTACAAACAGAAAAATTAACAGATTTGCTTTGGCTTTGCTAAATG 675
Db 64160 GAGTATGACTACAAACAGAAAAATTAACAGATTTGCTTTGGCTTTGCTAAATG 64101
Gy 676 TTACCTGACAAAATCTTAAGCCAGTTCTCAATTTTCGTTTGAATGAATACCTAGTTT 735
Db 64100 TTACCTGACAAAATCTTAAGCCAGTTCTCAATTTTCGTTTGAATGAATACCTAGTTT 64041
Gy 736 TAGTCAGAGGCTGGGCGGCAATAGCTGATGCTGTGATCCAGTCTTTGGCGGCCAG 795
Db 64040 TAGTCAGAGGCTGGGCGGCAATAGCTGATGCTGTGATCCAGTCTTTGGCGGCCAG 63981
Gy 796 GCAGGTGATCACTTAAGGTGAGAGTTTGAGACCAGCTTGCCCAACATGTGAAACGTT 855
Db 63980 GCAGGTGATCACTTAAGGTGAGAGTTTGAGACCAGCTTGCCCAACATGTGAAACGTT 63921
Gy 856 GTCTCTACTAAAAATTAACAAAATTAAGACAGGCGTGTGGGCAACATCTGTAAATCCAGCT 915
Db 63920 GTCTCTACTAAAAATTAACAAAATTAAGACAGGCGTGTGGGCAACATCTGTAAATCCAGCT 63861
Gy 916 ACTCAGAGGCTAACACAGAAAAATTTCTTAACCTGGAGAGCAGAGTTGACATGAGCC 975
Db 63860 ACTCAGAGGCTAACACAGAAAAATTTCTTAACCTGGAGAGCAGAGTTGACATGAGCC 63801
Gy 976 ATTGCATCCAGCTGTGGGCAACACAGTGAAGCTTGTCTCAAAAAAATTTTAAATCT 1035
Db 63800 ATTGCATCCAGCTGTGGGCAACACAGTGAAGCTTGTCTCAAAAAAATTTTAAATCT 63741

RESULT 3
AC104170/c 141670 bp DNA linear PRI 23-FEB-2002
LOCUS Homo sapiens chromosome 1 clone RP11-253A20, complete sequence.
DEFINITION AC104170 AL354990
ACCESSION AC104170.2 GI:18973866
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 141670)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
Direct Submission
Unpublished
2 (bases 1 to 141670)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (05-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 141670)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (23-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Feb 23, 2002 this sequence version replaced gi:17352434.

Genome Center

Center: University of Washington Genome Center

Center Code: UMG

Web site: <http://www.genome.washington.edu>Contact: uwgctgs@u.washington.edu

Drafting Center: SC

Project Information

Center project name: chr-1

Center clone name: RP11-253A20 (sc0660)

Summary Statistics

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator RT; 50% of reads

Chemistry: Dye-terminator Big Dye; 50% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 141542 bases at least Q40

Consensus quality: 141651 bases at least Q30

Consensus quality: 141670 bases at least Q20

Insert size: 141884; sum-of-contigs

Quality coverage: 8.5x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': RP5-932P4 (UMGC:sc0224) AL138904

3': Mapping in progress

Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:

all regions were either double-stranded or sequenced with an

alternate chemistry or covered by high quality data (i.e., Phred

quality >= 30); an attempt was made to resolve all sequencing

problems, such as compressions and repeats; all regions were

covered by at least one plasmid subclone or more than one M13

subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest

fingerprinting. Comparison of the experimentally derived digest

fragments with sequence-predicted fragments is given below.

The electronically-digested sequence consists of both insert and

vector, in order to accurately represent the entire circular BAC.

Small fragments below a variable cutoff (approximately 400-800 bp)

are not resolved in the fingerprint and hence do not appear

in the table. There are no significant remaining discrepancies

between the experimental and predicted values. Uniquely ordered

fragments are separated by dashed lines.

EcoRI HindIII BglII

SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt

	-----	-----	-----	-----
	1339	1373	5115	5021
	-----	-----	-----	-----
	768	<800	9156	9186
	-----	-----	-----	-----
	9943	9786	4300	4272
	-----	-----	-----	-----
	6286	6162	2420	2478
	-----	-----	-----	-----
	317	<800	3736	3657
	-----	-----	-----	-----
	2147	2184	732	739
	-----	-----	-----	-----
	1442	1373	-----	-----
	-----	-----	-----	-----
	3025	3095	-----	-----
	-----	-----	-----	-----
	871	891	-----	-----
	-----	-----	-----	-----
	1170	1162	-----	-----
	-----	-----	-----	-----
	2718	2641	-----	-----
	-----	-----	-----	-----
	1173	1162	-----	-----
FEATURES	Location/Qualifiers			
source	1..141670			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
	/chromosome="1"			
	/clone="RP11-253A20"			
	/clone_11b="RP11 human BAC library 11"			
	45..212			
ORIGIN	/note="Single subclone region"			
Query Match	97.5%	Score 101.6	DB 9	Length 141670;
Best Local Similarity	99.1%	Pred. No. 3.5e-187;		
Matches 1011	Conservative 5;	Mismatches 4;	Indels 0;	Gaps 0;
Qy	16	TTAATGTAATAAATTTCTATATGAATTTAATGGAAATTAAGCATCATAGAAAAA	75	
Db	20783	TTAATGTAATAAATTTCTATATGAATTTAATGGAAATTAAGCATCATAGAAAAA	20724	
Qy	76	TGCTCTACTGTTGAAAACATTATTGTTACATTTGGTCAACTATCTTCAATTACTT	135	
Db	20723	TGCTCTACTGTTGAAAACATTATTGTTACATTTGGTCAACTATCTTCAATTACTT	20664	
Qy	136	TTAGTACTATATGTTAAGTTGTAACGAGGCGAGCTTATATGTAAATGCGAGCTGAC	195	
Db	20663	TTAGTACTATATGTTAAGTTGTAACGAGGCGAGCTTATATGTAAATGCGAGCTGAC	20604	
Qy	196	AGCATGAATAATACATATCTATATTTTGTACTATCTTATAGGAAATCAGAGAAATT	255	
Db	20603	AGCATGAATAATACATATCTATATTTTGTACTATCTTATAGGAAATCAGAGAAATT	20544	
Qy	256	CAAAACCTGTTAGTTTAAAGGTAATGTCACATTTTAAATGTGCGTATATTATATAC	315	
Db	20543	CAAAACCTGTTAGTTTAAAGGTAATGTCACATTTTAAATGTGCGTATATTATATAC	20484	
Qy	316	ATGATTTGACGTTGTGAAAATATTTTCCCGAGCTTTTATTTAGATGAGATCTACAGT	375	
Db	20483	ATGATTTGACGTTGTGAAAATATTTTCCCGAGCTTTTATTTAGATGAGATCTACAGT	20424	
Qy	376	GTAAGCAAACTATATATCTGTCAACTCATAGTGCATAGTCAGACTGATCCCATG	435	
Db	20423	GTAAGCAAACTATATATCTGTCAACTCATAGTGCATAGTCAGACTGATCCCATG	20364	
Qy	436	CTAAAAATATATGTTGTTAAATAGCTTTGTTAAATGTTGTTAGTCACTATACCA	495	
Db	20363	CTAAAAATATATGTTGTTAAATAGCTTTGTTAAATGTTGTTAGTCACTATACCA	20304	

Qy	496	AGCTTCAGAGKATTATACATTATAAAAACTTGKTTTATCTTGGAATACGTTTTT	555
Db	20303	AGCTTCAGAGTATTACATTATATAAAACCTTGTTTTTATCTTGGAATACGTTTTT	20244
Qy	556	TCGATGCAAAAGTTAAATTTCTTCAGCCTTAATTTTTTATTAATATATAGATGTGAT	615
Db	20243	TCGATGCAAAAGTTAAATTTCTTCAGCCTTAATTTTTTATTAATATATAGATGTGAT	20184
Qy	616	GAGTATGACTACAAAACAGGAAAAATTAACAGATTTGTTTGCTTTGCTTAATTG	675
Db	20183	GAGTATGACTACAAAACAGGAAAAATTAACAGATTTGTTTGCTTTGCTTAATTG	20124
Qy	676	TTACCTGCAAAAATCTTAGCCAGTTCTTCATTTTCTGTTTGAGATGAAGTACTAGTTT	735
Db	20123	TTACCTGCAAAAATCTTAGCCAGTTCTTCATTTTCTGTTTGAGATGAAGTACTAGTTT	20064
Qy	736	TAGTCAGAGGGCTGGGCGCATAGCTGATGCTGTGTGCTCCAGTCTTTGCGGGCCGAG	795
Db	20063	TAGTCAGAGGGCTGGGCGCATAGCTGATGCTGTGTGCTCCAGTACTTTGCGGGCCGAG	20004
Qy	796	GCAGGTGATCATCTTAAGTATAGAGATTGAGACCAAGCTGCCAATGTTGAACCTT	855
Db	20003	GCAGGTGATCATCTTAAGTATAGAGATTGAGACCAAGCTGCCAATGTTGAACCTT	19944
Qy	856	GTCCTACTAAAAATACAAAATTAGACAGGCGTGTGTCGACACATCTGTATTCCAGCT	915
Db	19943	GTCCTACTAAAAATACAAAATTAGACAGGCGTGTGTCGACACATCTGTATTCCAGCT	19884
Qy	916	ACTCAGAGGCTTAACACAGGAAAAATTCCTTGAACCTGGAGGACAGAGTTGCACTGAGCC	975
Db	19883	ACTCAGAGGCTTAACACAGGAAAAATTCCTTGAACCTGGAGGACAGAGTTGCACTGAGCC	19824
Qy	976	ATTGCACTCCAGCCTGGGCAACACAGTGAAGCTCTTGTCTCAAAAAAATTTTAACT	1035
Db	19823	ATTGCACTCCAGCCTGGGCAACACAGTGAAGCTCTTGTCTCAAAAAAATTTTAACT	19764
RESULT 4			
LOCUS	ALJ59698	164500 bp	DNA linear HTG 10-JUL-2001
DEFINITION	Homo sapiens chromosome X clone RP11-55316, 3 unordered pieces.		
ACCESSION	ALJ59698		
VERSION	ALJ59698.10	GI:9930965	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_CANCELLED.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 Kimberley, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk		
COMMENT	requests: clonerequests@sanger.ac.uk On Aug 27, 2000 this sequence version replaced gi:9717116. ----- Genome Center Center: Sanger Centre Center code: SC Web site: http://www.sanger.ac.uk Contact: humquery@sanger.ac.uk ----- Project Information Center project name: ba55316 ----- Summary Statistics Assembly program: XGAP4; version 4.5 Sequencing vector: plasmid; L08752; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Consensus quality: 163671 bases at least Q40 Consensus quality: 163958 bases at least Q30 Consensus quality: 164134 bases at least Q20 Insert size: 164300; sum-of-contigs Insert size: 175514; agarose-fp Quality coverage: 9.44x in Q20 bases; sum-of-contigs Quality coverage: 8.89x in Q20 bases; agarose-fp		

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* As soon as it is available and the accession number will
* be preserved.
*
* 1      23821: contig of 23821 bp in length
*
* 23822      gap of 100 bp
*
* 23922      31390: contig of 7469 bp in length
*
* 31391      31490: gap of 100 bp
*
* 31491      164500: contig of 133010 bp in length.
*
* Location/Qualifiers

```

```

source
1. .164500
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-55316"
/clone_1fb="RPC1-11.2"
1. .23821
/note="assembly_fragment:02323
clone_end:SP6
vector_side:left"
23922. .31390
/note="assembly_fragment:01611"
31491. .164500
/note="assembly_fragment:01777"

```

Query Match	97.5%	Score 1011.6	DB 2	Length 164500
Best Local Similarity	99.1%	Pred. No. 3.5e-187		
Matches 1011	Conservative	5	Mismatches 4	Indels 0
			Gaps	0
QY	16	TTAATGTATAAAATATTTCTTATATGAAATTTTATGAGGAATTTAGACATCATGAAAAA	75	
DB	56406	TTAATGTATATAAATATTTCTTATATGATTTTAAATGGAAATTAGACATCATGAAAAA	56465	
QY	76	TGCTCTTACGTTGAAAAACCTTATTTGTTACATTTTGGTCACTAATCTTTCATTA	135	
DB	56466	TGCTCTTACGTTGAAAAACCTTATTTGTTACATTTTGGTCACTAATCTTTCATTA	56525	
QY	136	TTAGTAATATATATGTTAAGTTGTACAGGCGAGCTTATATAGTAAATGCGACGTGAC	195	
DB	56526	TTAGTAATCATATATGTTAAGTTGTACAGGCGAGCTTATATAGTAAATGCGACGTGAC	56585	
QY	196	AGATGAAAAATTAACATATCTTAATATTTTGTGACTATCTTATTAGGAAATCAGAAATTT	255	
DB	56586	AGATGAAAAATTAACATATCTTAATATTTTGTGACTATCTTATTAGGAAATCAGAAATTT	56645	
QY	256	CAAAACCTGTTAGTTTTAGGGTATAGCAATTTTATTAATGTGCGGTATATTATATAC	315	
DB	56646	CAAAACCTGTTAGTTTTAGGGTATAGCAATTTTATTAATGTGCGGTATATTATATAC	56705	
QY	316	ATGATTTGACGTTGTGMAAATATTTTCCGTGACCTTTATTTTATGATGAGATCTACAGT	375	
DB	56706	ATGATTTGACGTTGTGMAAATATTTTCCGTGACCTTTATTTTATGATGAGATCTACAGT	56765	
QY	376	GTAGGCAAACTATATATCTGTCAACCTCATTAAGTGTCAATGATGACATCCCATG	435	
DB	56766	GTAGGCAAACTATATATCTGTCAACCTCATTAAGTGTCAATGATGACATCCCATG	56825	
QY	436	CTAAATATTAATAGTGTAKAAATACGCTTTGTGAAATGTTGTGTAGTCAATTAACCA	495	
DB	56826	CTAAATATTAATAGTGTAKAAATACGCTTTGTGAAATGTTGTGTAGTCAATTAACCA	56885	
QY	496	AGTCTTCAAGKATTAATTAATAAAACCTTGKITTATATCTGTGAATCCGTTTTT	555	
DB	56886	AGTCTTCAAGKATTAATTAATAAAACCTTGKITTATATCTGTGAATCCGTTTTT	56945	
QY	556	TCCATGCAAAATTAATAATCTTCAACCCCTTAATTTTTTTTATTAATATATAGATGTGAT	615	

Db 56946 TCGATGCAAGTTAAATAATCTTCAGCCCTTTAATTTTTTTATATATATAGATGTGAT 57005

QY 616 GAGTATGACTACAAAACAGAAAAAATTAACAGATTTCGTTGAGCTTTGGTAAATG 675

Db 57006 GAGTATGACTACAAAACAGAAAAAATTAACAGATTTCGTTGAGCTTTGGTAAATG 57065

QY 676 TTACTGCAAAAATCTTAGCCAGTTCCTCAATTTTCGTTTGAGATGAAGATCTTAGTT 735

Db 57066 TTACTGCAAAAATCTTAGCCAGTTCCTCAATTTTCGTTTGAGATGAAGATCTTAGTT 57125

QY 736 TAGTCACAGAGGCTGGGGCGGATAGCTGTAATGCCCTGTGCTCCAGTCTTTGCGGGGCGAG 795

Db 57126 TAGTCACAGAGGCTGGGGCGGATAGCTGTAATGCCCTGTGATCCCAATGACTTTGGCGGGGCGAG 57185

QY 796 GCAGGTGATCACTTAAAGTTCAGAGATTGTAGACACAGCTGCACCACATGATGTAACGTT 855

Db 57186 GCAGGTGATCACTTAAAGTTCAGAGATTGTAGACACAGCTGCACCACATGATGTAACGTT 57245

QY 856 GTCTCTACTTAAAAATTAAGAAAATTTAGACAGGCGTGTGGCACACATCTGTAAATTCAGCT 915

Db 57246 GTCTCTACTTAAAAATTAAGAAAATTTAGACAGGCGTGTGGCACACATCTGTAAATTCAGCT 57305

QY 916 ACTAGAGAGGCTAACACAGGAAAATTCCTTGAACCTGGAGGCAAGAGTTGCAGTGAGCC 975

Db 57306 ACTAGAGAGGCTAACACAGGAAAATTCCTTGAACCTGGAGGCAAGAGTTGCAGTGAGCC 57365

QY 976 ATTGCACTCCAGGCTGGGGCAACACAGTGAAGACTTTGTCTCAAAAAAAAAAAAAAACT 1035

Db 57366 ATTGCACTCCAGGCTGGGGCAACACAGTGAAGACTTTGTCTCAAAAAAAAAAAAAAACT 57425

RESULT 5	AL731898/c	169304 bp	DNA	linear	HTG 12-MAY-2002
LOCUS	AL731898				
DEFINITION	Homo sapiens chromosome 1 clone RP11-449B11.				
ACCESSION	AL731898				
VERSION	AL731898.3	GI:20795667			
KEYWORDS	HTG; HTGS_PHASE2; HTGS_CANCELLED.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
	1 (base 1 to 169304)				
REFERENCE	Harrison,E.				
AUTHORS	Direct Submission				
TITLE	Submitted (08-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,				
JOURNAL	CambridgeShore, CB10 1SA, UK. E-mail enghires: humphreys@sanger.ac.uk				
COMMENT	On May 14, 2002 this sequence version replaced gi:20429020.				

```

Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
Project Information
Center project name: Bk449B11
-----
Summary Statistics
Assembly program: XGAP4; version: 4.5
Sequencing vector: M13; M77815; 54% of reads
Sequencing vector: plasmid; U08752; 45% of reads
Chemistry: dye-terminator Big Dye; 99% of reads
Chemistry: dye-terminator; 0% of reads
Consensus quality: 169292 bases at least Q40
Consensus quality: 169299 bases at least Q30
Consensus quality: 169301 bases at least Q20
Insert size: 169304; sum-of-contigs
Insert size: 178813; 2.3% error; agarose-fp
Quality coverage: 18.0% in Q20 bases; sum-of-contigs
coverage: 17.90% in Q20 bases; agarose-fp

```

* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs


```
repeat_region      383..680
                    /rpt_family="AluSc"
misc_feature        /note="DPS similarity to AA229025 nc50c11.s1 NCI CGAP_Pr3
                    Homo sapiens cDNA clone IMAGE:1011572. Score: 550
                    identity: 290/300 (96%)."
```

```
repeat_region      1281..1318
                    /rpt_family="POLY_A"
repeat_region      complement(1443..1755)
                    /rpt_family="AluSp"
repeat_region      1917..2204
                    /rpt_family="AluSc"
repeat_region      complement(2735..2777)
                    /rpt_family="(CA)n"
repeat_region      3566..3652
                    /rpt_family="FRAM/FAM"
repeat_region      3671..3719
                    /rpt_family="MER4D"
repeat_region      complement(3720..4019)
                    /rpt_family="AluSp"
repeat_region      4047..4324
                    /rpt_family="MER4D"
misc_feature        4960..5217
                    /note="DPS similarity to multiple ESTs:
                    (4960..5217) AA446441 zw60d10.r1 Soares total fetus NB2HP8
                    9w Homo sapiens cDNA clone 774451 5' similar to contains
                    LTR3.t2 LTR3 repetitive element ;
                    (1..259) ; 98% identity.~(4962..5216) AA195132 zr34b08.r1
                    Soares NbHMPu S1 Homo sapiens cDNA clone 665271 5' similar
                    to contains LTR3.b1 LTR3 repetitive element; (1..253) ; 99%
                    identity.~(4962..5216) W03459 za06e09.r1 Soares melanocyte
                    2NbHM Homo sapiens cDNA clone 291784 5' similar to
                    contains LTR3.b1 LTR3 repetitive element;(1..253) ; 96%
                    identity.
                    and others.~".
                    complement(5599..5875)
                    /rpt_family="LINE2"
repeat_region      complement(6952..6980)
                    /rpt_family="AT_rich"
repeat_region      7067..7364
                    /rpt_family="AluSg"
repeat_region      complement(7488..7593)
                    /rpt_family="MR"
repeat_region      7718..8134
                    /rpt_family="LTR3"
repeat_region      complement(8649..9102)
                    /rpt_family="LINE2"
repeat_region      9114..9386
                    /rpt_family="AluSc"
misc_feature        9573..9676
                    /note="predicted exon, program: Gra12exons_human_1.3,
                    frame: 1, quality: excellent, score: 90.000"
repeat_region      10469..10758
                    /rpt_family="AluSg"
repeat_region      11041..11377
                    /rpt_family="MER74"
misc_feature        11448..11576
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (1..43) ; 48% identity.-Other
                    overlapping matches:
                    (11508..11634) predicted exon, program:
                    Gra12exons_human_1.3, frame: 2, quality: excellent,
                    score: 83.000"
misc_feature        12108..12167
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (45..63) ; 26% identity."
```

```
repeat_region      12861..12900
                    /rpt_family="MER4D"
misc_feature        12973..13537
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (64..249).-Other overlapping
                    matches:
                    (12973..15108) predicted exon, program:
```

```
CDS
                    Gra12exons_human_1.3, frame: 0, quality: excellent,
                    score: 87.000"
                    13033..15108
                    /note="hypothetical ZNF-like protein"
                    /codon_start=1
                    /product="R30217.1"
                    /protein_id="AA57932.1"
                    /db_xref="GI:2822143"
                    /translation="MCSIIKLILHLAHDGTHPEOGLYTCAAEHLHQEKQIREKLT
                    RSDWRPSFVNSAHVGERNFCTOGKDPFTASDILQOQVLNSGKMLYDTODGEAF
                    QGEONDPNSSOGKDPFCHQGLFEHOKTHGERPRPEEGELFRYNSNLITKQONHA
                    GERPEEGTEYKFTIRKSNLVQHOKIHSEGLSKRSPDIEHOLLSRPPECTOGCK
                    AFLTOAHLVGHOKTHGEOPPECKCKPFMVNSKILRHOKYVTGERRRSCGGLF
                    MDSPTLGRHRYHTGERPPECSCICGFRFHSRSTLNMHQRVHAGKRLKSCSGKAVSL
                    KANVQHLKHTGSRPECTECERKAFVRKSHLVHOKIHTDASKSDLIQNRIDIR
                    PRPYTSCGCKAFLTOAHLVGHOKIHTGERPYCTOCANAFVKSHLVHOKIHTDAF
                    SKRSDLIQNRIDIRPRPYTSCGCKAFLTOAHLVGHOKIHTGERPYCTOCANAFVKSHLVHOKIHTDAF
                    SYKLVIHQRIHTGEKPYKSCGKFRVRCITLSRHOKYHTGERPYKSCGKFRPDSY
                    KLIIHQVHTGEKPYKSCGKFRVRCITLSRHOKYHTGERPYKSCGKFRPDSY
                    LIHQSHRGSPPKLABCGKDPFKCNTGQRQKTHGERRSTECGSSKVRKNSLTIKH
                    QIIHTGRRP"
                    13592..14095
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (250..417)."
```

```
misc_feature        14147..14314
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (418..473)."
```

```
misc_feature        14366..15116
                    /note="DPS similarity to (U66561) kruppel-related zinc
                    finger protein [Homo sapiens] (474..726)."
```

```
repeat_region      15272..15565
                    /rpt_family="AluSc"
repeat_region      complement(15566..15720)
                    /rpt_family="LIMB3"
repeat_region      15721..16037
                    /rpt_family="MER1B"
repeat_region      complement(16055..16255)
                    /rpt_family="LIMB3"
repeat_region      16256..17064
                    /rpt_family="MER7B"
repeat_region      complement(17065..17363)
                    /rpt_family="AluSg"
repeat_region      complement(17364..17658)
                    /rpt_family="AluSg"
repeat_region      17659..17849
                    /rpt_family="MER7B"
repeat_region      complement(17853..18015)
                    /rpt_family="LIMB3"
repeat_region      18024..18050
                    /rpt_family="(TAA)n"
repeat_region      complement(18051..18348)
                    /rpt_family="AluY"
repeat_region      complement(18358..19222)
                    /rpt_family="LIMB3"
misc_feature        19980..20439
                    /note="DPS similarity to N20479 yx39g09.s1 Homo sapiens
                    cDNA clone 264160 3'. Score: 900 Identity: 455/460
                    (98%)."
```

```
repeat_region      20474..20521
                    /rpt_family="POLY_A"
repeat_region      complement(20832..20969)
                    /rpt_family="MSTC"
repeat_region      complement(21068..21397)
                    /rpt_family="AluSg"
repeat_region      complement(21402..21590)
                    /rpt_family="MSTC"
repeat_region      21846..22198
                    /rpt_family="MSTC"
repeat_region      complement(22532..22641)
                    /rpt_family="L1MA9"
repeat_region      complement(22738..22856)
                    /rpt_family="(GGA)n"
repeat_region      complement(22871..23168)
```


gene	/rpt_family="MER58A" 7244. .7983 /gene="TCRAV18S1" /pseudo 7244. .7263 /gene="TCRAV18S1" /standard_name="TCRAV18S1" /pseudo join(<7482. .7527,7656. .>7944) /gene="TCRAV18S1" /product="TCRAV18S1" /pseudo join(7482. .7527,7656. .>7944) /gene="TCRAV18S1" /standard_name="TCRAV18S1" /pseudo 7945. .7951 /gene="TCRAV18S1" /standard_name="TCRAV18S1" /note="heptamer" 7952. .7974 /gene="TCRAV18S1" /standard_name="TCRAV18S1" /note="spacer" 7975. .7983 /gene="TCRAV18S1" /standard_name="TCRAV18S1" /note="nonamer" complement(8236. .8389) /rpt_family="MER69A" complement(8556. .8740) /rpt_family="L1MB7" complement(9281. .10170) /rpt_family="L1PA2" complement(10023. .11230) /rpt_family="L1" 11234. .11405 /rpt_family="L1" 11818. .12491 /gene="TCRAV19S1" 11818. .11837 /gene="TCRAV19S1" /standard_name="TCRAV19S1" /note="Old name TCRAV12S1" join(<11900. .11948,12152. .>12452) /gene="TCRAV19S1" /product="TCRAV19S1" /note="Old name TCRAV12S1" join(11900. .11948,12152. .>12452) /gene="TCRAV19S1" /standard_name="TCRAV19S1" /note="Old name TCRAV12S1" /codon_start=1 /product="TCRAV19S1" /protein_id="AAB69019.1" /db_xref="GI:2358044" /translation="MLTASILRVAISICVSSMSM VERPDITVLYFYKOPSGSELVFLIRNSFDQNN QVDSAVYCAISER" join(11900. .11948,12152. .>12452) /gene="TCRAV19S1" /standard_name="TCRAV19S1" /note="Old name TCRAV12S1" 12453. .12459 /gene="TCRAV19S1" /standard_name="TCRAV19S1" /note="heptamer" 12460. .12482 /gene="TCRAV19S1" /standard_name="TCRAV19S1" /note="spacer" 12483. .12491 /gene="TCRAV19S1"
promoter	
mRNA	
CDS	
V_segment	
misc_recomb	
misc_recomb	
misc_recomb	

```

repeat_region      /standard_name="TCRAVL9S1"
                    /note="nonamer"
                    13763. .1833
repeat_region      /rpt_family="MIR2"
                    13879. .13946
repeat_region      /rpt_family="MIR"
                    complement(19382. .19677)
repeat_region      /rpt_family="AluSg"
                    21467. .21622
repeat_region      /rpt_family="MIR"
                    23034. .23706
repeat_region      /rpt_family="L1ME3A"
                    23955. .24354
repeat_region      /rpt_family="MSTA"

```

Query Match	19.9%;	Score 206.2;	DB 9;	Length 251124;
Best Local Similarity	81.8%;	Pred. No. 1,6e-30;		
Matches 238;	Conservative	0;	Mismatches 53;	Indels 0; Gaps 0;
OY	743	GGGCGTCGGCCGGAGTACGTATGCCTGTGTGCCAGTGCTTTGGCGGGGCGCAGGCAAGTGTG	802	
Db	114854	GAGGCTGGGCACATTGGCTCATGCTCTGTATATCCACACTTTTGGAGGCTGAGGCGGGCG		
OY	803	GATCACTTAAGGTCAGAGATTGTAGACCAGCCTGCCCAACATGTTGAAAAGTTGTCTTA	862	
Db	114914	AATCACCTGAGGTCAGAGATTGTAGAGACTAGCCTGGCCAATGTGTAAACCCGTGCTTTA	114973	TATTCACAGCTACTCCAGG
OY	863	CTAAAAATATCAAAAATTTAGACAGCGCGTGTGTGCACACATCTGTATATTCACAGCTACTCCAGG	922	
Db	114974	CTAAAAATATCAAAAATTTAGCCAGGCAATGTGTATGACACACTGTATATCAACAGCTACTCCGGG	115033	BCTATACAGCTACTCCGGG
OY	923	AGGCTAACACACAGAAAAATTCCTTGAACCTGGGAGGAGAGTTGSCAGTAGGCCATTGCAC	982	
Db	115034	AGGCTGAGGCAAGAGAAATCGCTTGAACCTGGAGGTGGAGGTTGCTGTAGACCACTGCAC	115033	AGGCTGAGGCAAGAGAAATCGCTTGAACCTGGAGGTGGAGGTTGCTGTAGACCACTGCAC
OY	983	TCCAGCCTGGGCAACACAGTAGAGACTCTTGTCTCAAAAAAAAAAAAAAAAAA	1033	
Db	115094	TCCAGCCTGGGCAACACAGTAGAGACTCTGTCTCAAAAAAAAAAAAAAAAAA	115144	
RESULT 8				
AC008666		99108 bp	DNA	linear PRI 28-FEB-2001
LOCUS	Homo sapiens chromosome 5 clone CTR-35A8,			complete sequence.
DEFINITION	AC008666			
ACCESSION	AC008666.5	GI:13162497		
VERSION				
KEYWORDS	HTG.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 99108)			
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.			
JOURNAL	Direct Submission			
TITLE	Unpublished			
REFERENCE	2 (bases 1 to 99108)			
AUTHORS	DOE Joint Genome Institute.			
JOURNAL	Direct Submission			
TITLE	Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA 3 (bases 1 to 99108)			
REFERENCE	DOE Joint Genome Institute and Stanford Human Genome Center.			
AUTHORS	Submitted (28-FEB-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA On Feb 28, 2001 this sequence version replaced gi:7709249.			
JOURNAL	Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov			
COMMENT	Finishing Completed at Stanford Human Genome Center www.hgc.stanford.edu Quality: Phrap Quality >=40 99.6% of Sequence; Estimated Total Number of Errors is 0.3.			
	STS Content:			

SHGC-10972 G11217
WI-5180 G02907.
Location/Qualifiers
1..99108
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTB-35A8"

ORIGIN

Query Match 19.8%; Score 205.4; DB 9; Length 99108;
Best Local Similarity 79.5%; Pred. No. 2.5e-30;
Matches 260; Conservative 0; Mismatches 56; Indels 11; Gaps 1;

Qy 718 GATGAAGTACTTACTGTTTAACTCCAGGGGCTGGCGCATAGCTGATGCTGTGTGCCA 777
Db 51492 GAGCAGAGAAACAAGTTGCTCCCTGGGCTAGGTGTGTGCTCATGCTGTATATCCA 51551
Qy 778 GTGCTTGGCGGGCGGAGGAGGTGATCACTTAAGTCAGAGATTGAGACCAAGCCCTGC 837
Db 51552 GCACCTTGGGAGACTGAGGCGAGGTGATCACTTGAGTCAAGAGTTTGATACCAAGCTTG 51611
Qy 838 CCAACATGTGTAACGTTGTCTCTACTAATAAATAAATAATTAGACAGCGCTGTGGCAC 897
Db 51612 CCAACATGTGTAACCCCGTCTGTACTGAATAAATAAATAATTAGCAGGCGATGTGGCAC 51671
Qy 898 ACATCTGTAATCCAGCTACTCAGAGGCTTAACACAGAAATTCCTTGAACCTGGAGG 957
Db 51672 ACACCTGTAGTTCACAGCTACTCAGAGGCTGAGGACAGATCTCTTGAACCTGGAGG 51731
Qy 958 CAGAGGTTGCAGTGA-----GCCATTGACCTCCAGCCTGGGCAACAGAGTGGA 1006
Db 51732 TGGAGGTTCAGTGAAGCTGAGATCATGCTACTGAGTCAGCCTGGGCAACAGAGTGGA 51791
Qy 1007 CTCTGTCTCAAAAAAAAAAAAAAAAAA 1033
Db 51792 CTCTGTCTTAAAAAAAAAAAAAAAAAA 51818

RESULT 9
AC008641/c 176629 bp DNA linear PRI 23-AUG-2001
LOCUS Homo sapiens chromosome 5 clone CTB-17P3, complete sequence.
AC008641
AC008641.7 GI:15281183
VERSION
KEYWORDS
HTG.
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
DOE Joint Genome Institute and Stanford Human Genome Center.
DIRECT SUBMISSION
Unpublished
2 (bases 1 to 176629)
DOE Joint Genome Institute.
REFERENCE
AUTHORS
TITLE
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 176629)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
AUTHORS
TITLE
Direct Submission
Submitted (21-OCT-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 176629)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
AUTHORS
TITLE
Direct Submission
Submitted (23-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Aug 23, 2001 this sequence version replaced gi:10944452.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >40.99.5% of Sequence;
Estimated Total Number of Errors is 0.8.
STS Content:
WI-5180 G02907
SHGC-10972 G11217
WI-12039 G22831.
Location/Qualifiers
1..176629
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTB-17P3"

ORIGIN

Query Match 19.8%; Score 205.4; DB 9; Length 176629;
Best Local Similarity 79.5%; Pred. No. 2.4e-30;
Matches 260; Conservative 0; Mismatches 56; Indels 11; Gaps 1;

Qy 718 GATGAAGTACTTACTGTTTAACTCCAGGGGCTGGCGCATAGCTGATGCTGTGTGCCA 777
Db 47412 GAGCAGAGAAACAAGTTGCTCCCTGGGCTAGGTGTGTGCTCATGCTGTATATCCA 47353
Qy 778 GTGCTTGGCGGGCGGAGGAGGTGATCACTTAAGTCAGAGATTGAGACCAAGCCCTGC 837
Db 47352 GCACCTTGGGAGACTGAGGCGAGGTGATCACTTGAGTCAAGAGTTTGATACCAAGCTTG 47293
Qy 838 CCAACATGTGTAACGTTGTCTCTACTAATAAATAAATAATTAGACAGCGCTGTGGCAC 897
Db 47292 CCAACATGTGTAACCCCGTCTGTACTGAATAAATAAATAATTAGCAGGCGATGTGGCAC 47233
Qy 898 ACATCTGTAATCCAGCTACTCAGAGGCTTAACACAGAAATTCCTTGAACCTGGAGG 957
Db 47232 ACACCTGTAGTTCACAGCTACTCAGAGGCTGAGGACAGATCTCTTGAACCTGGAGG 47173
Qy 958 CAGAGGTTGCAGTGA-----GCCATTGACCTCCAGCCTGGGCAACAGAGTGGA 1006
Db 47172 TGGAGGTTCAGTGAAGCTGAGATCATGCTACTGAGTCAGCCTGGGCAACAGAGTGGA 47113
Qy 1007 CTCTGTCTCAAAAAAAAAAAAAAAAAA 1033
Db 47112 CTCTGTCTTAAAAAAAAAAAAAAAAAA 47086

RESULT 10
AC106805/c 205819 bp DNA linear HTG 25-JAN-2002
LOCUS Homo sapiens chromosome 5 clone RP11-507E2, WORKING DRAFT SEQUENCE.
AC106805
AC106805.1 GI:18139355
VERSION
KEYWORDS
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
DOE Joint Genome Institute.
DIRECT SUBMISSION
Unpublished
2 (bases 1 to 205819)
DOE Joint Genome Institute.
REFERENCE
AUTHORS
TITLE
Direct Submission
Submitted (12-JAN-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information

Db	7006	ATGCGGCACCTGCACCTGCAGCTCGGACAGACAGCCACTCTGTGTTTCAAAAAAAAAA	7011
OY	1028	AAAAA 1033	
Db	70120	AAAAA 70125	
RESULT 12			
ACLI0594/c			
LOCUS	ACLI0594	159490 bp	DNA linear PRI 13-SEP-2002
DEFINITION	Homo sapiens chromosome 17, clone CTD-2193J24, complete sequence.		
ACCESSION	ACLI0594		
VERSION	ACLI0594.5	GI:22830329	
KEYWORDS	HTG		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Bitren,B., Nusbaum,C. and Lander,E.		
TITLE	Homo sapiens chromosome 17, clone CTD-2193J24		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 159490)		
TITLE	Bitren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Baatien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campiolo,A., Chang,J., Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., Lahocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C., McCarthy,M., Mcgwan,P., McKernan,K., Meldrum,J., Menus,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Notman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tophan,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,D., Zembek,L., Zimmer,A. and Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 159490)		
AUTHORS	Bitren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Baatien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrum,J., Menus,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Notman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Tophan,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	4 (bases 1 to 159490)		
AUTHORS	Bitren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Baatien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,		

Camara, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dekrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, J., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karats, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Plunkhamp, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnpack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Thodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, W.

Direct Submission
Submitted (13-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 330 Charles Street, Cambridge, MA 02141, USA
On Sep 13, 2002 this sequence version replaced gi:2296745.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L25381
Center clone name: 2193_J_24

Only the first 159,5 kilobases of this clone are being submitted.
The remainder overlaps accession number AC003042 [WICGR project L260].

FEATURES
source
Location/Qualifiers
1.159490
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="CTD-2193J24"
/clone_lib="CTDI Human BAC"
complement(5095..5617)
/rpt_family="1LMD3"
8205..8396
/rpt_family="AluSg/X"
8399..8482
/rpt_family="L1R65"
8505..8633
/rpt_family="F1L1M A"
complement(8643..8779)
/rpt_family="AluJo"
8780..8996
/rpt_family="L1R65"
complement(8998..9076)
/rpt_family="MIR"
complement(10090..10879)
/rpt_family="HAI1"
complement(10880..10989)
/rpt_family="F1L1M C"
complement(10990..11042)
/rpt_family="HAI1"
complement(11056..11357)
/rpt_family="AluSx"
complement(11367..11430)
/rpt_family="HAI1"
12628..13088
/rpt_family="AluSx"
15961..16045
/rpt_family="Ga-rich"

```
repeat_region      16215. .16446  
/rpt_family="ALuNb"  
repeat_region      complement(16770. .17108)  
/rpt_family="ALuNb"  
repeat_region      18252. .18570  
/rpt_family="ALuNb"  
repeat_region      complement(18770. .18989)  
/rpt_family="H1R"  
unare             complement(18897. .19024)  
/note="single clone coverage"  
repeat_region      complement(19056. .19189)  
/rpt_family="LTR28"  
unare             complement(19236. .19253)  
/note="<30 qual SNGL region"  
repeat_region      complement(19244. .20030)  
/rpt_family="LTR1B"  
unare             19818. .20633  
/note="PCR product sequence only"  
repeat_region      complement(20031. .20388)  
/rpt_family="HUKRS-P2"  
repeat_region      complement(20468. .21174)  
/rpt_family="HUKRS-P2"  
repeat_region      complement(221174. .24069)  
/rpt_family="HUKRS-P2"  
repeat_region      24070. .24364  
/rpt_family="ALuSg"  
repeat_region      complement(24365. .25300)  
/rpt_family="HUKRS-P2"  
repeat_region      complement(25436. .25569)  
/rpt_family="LTR28"  
repeat_region      complement(25594. .26466)  
/rpt_family="LTR1B"  
repeat_region      complement(26936. .27052)  
/rpt_family="L1ME1"  
repeat_region      complement(27053. .27348)  
/rpt_family="ALuSc"  
repeat_region      complement(27349. .28223)  
/rpt_family="L1ME1"  
repeat_region      complement(28259. .28436)  
/rpt_family="ALuNb"  
repeat_region      complement(28442. .28555)  
/rpt_family="L1ME1"  
repeat_region      28556. .28643  
/rpt_family="FLAM_A"  
repeat_region      complement(28644. .28949)  
/rpt_family="ALuSx"  
repeat_region      28950. .28994  
/rpt_family="FLAM_A"  
repeat_region      complement(29006. .29119)  
/rpt_family="L1ME1"  
repeat_region      complement(29120. .29418)  
/rpt_family="ALuY"  
repeat_region      complement(29419. .29571)  
/rpt_family="L1ME1"  
repeat_region      complement(29583. .29744)  
/rpt_family="L1ME1"  
repeat_region      complement(29794. .29997)  
/rpt_family="MER20"  
repeat_region      complement(30024. .30125)  
/rpt_family="L1ME1"  
repeat_region      complement(30126. .30263)  
/rpt_family="FLAM_A"  
repeat_region      complement(30264. .30551)
```

	Query	Match	Score	DB	Length	159490
	Best Local Similarity	81.1%	Pred. No. 7e-30			
	Matches	236	Conservative	0	Mismatches	55
					Indels	0
					Gaps	0
Qy	743	GGGGCTGGGGCCAGTATGCTGTGGTCCAGTCTTTGGGGCCAGGACAGTG	802			
Db	86873	GGGGCCAGGCTGTGGCTCACTCTGTAACTTTCAGCATTTGGAGGGCCAGGTGGAG	8681			
Yy	803	GATCACTTAAAGTCAGGAGTTTGAACAAGCTTCCCAACATGTGTGAAGAAGCTTCTCTA	862			

Db	86613	GATCACCCTGGAGCTCAGGAGTTCCAGACCCAGCCTGGTCAACATGCTAAAGCCCGTCTTA	8675
Qy	863	CTAAAAATACAAAATTTAGACAGCCGTGGTGACACATCTGTAATTCCAGTACTCAG	922
Db	86753	CTAAAAATACAAAATTTAGCCAGGCGGTGATGCAACCCGCTGTAACTCCAGTACTTTAG	8669
Qy	923	AGGCTAACAGGAAATTCCTTGAACCTTGAGAGGACAGTTCAGTGGAGCATTGCAC	982
Db	86693	AGGCTGAGGAGGAGAACTCGGTTGAACCGGGGAAGCAGAGTTGGAGTGAGCCATTGCAC	8663
Qy	983	TCCAGCCTGGGCAACACAGTGAAGCTCTTGTCTCAAAAAAAAAAAAAAAAAA	1033
Db	86633	TCCAGCCTGGGCAACAGATGAGATCTTCATCTCAAAAAAAAAAAAAAAAAA	86583

RESULT_13	AC108026/c	AC108026	123902 bp	DNA	linear	PRI 13-MAR-2007
LOCUS		Homo sapiens	BAC clone Rpl1-1E15	from	4,	complete sequence.
DEFINITION		AC108026				
ACCESSION		AC108026.5	GI:19387784			
VERSION						
KEYWORDS		HTG.				
SOURCE		Homo sapiens	(human)			
ORGANISM		Homo sapiens				

REFERENCE
AUTHORS
TITLE
JOURNAL
1 (bases 1 to 123902)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8(11), 1097-1108 (1998)

REFERENCE 2 (bases 1 to 123902)

JOURNAL
REFERENCE
3 (bases 1 to 123902)
Unpublished (2001)

JOURNAL Submitted (24-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis MO 63108, USA

REFERENCE 4 (bases 1 to 123502)

TITLE	Direct Submission
JOURNAL	Submitted (23-FEB-2002) Genome Sequencing Center, Washington

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
5 (bases 1 to 123902)		Direct Submission		
	Waterston, R.	Submitted (13-MAR-2002)	Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA	On Mar 13, 2002 this sequence version replaced gl:1887888.

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watsen.wustl.edu
 ----- Summary Statistics -----
 Center project name: H_NH0041E15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatemio, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.edu> (PBCe3.6 VECTOR).

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-8111L19; the clone sequenced to the right is RP11-270L13, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-41B15; actual end is at base position 52485 of RP11-270L13.

Polymorphisms exists between AC108026, AC112724 and AC108043. Data from AC112724 was used to finish AC108026.

FEATURES

source

1..123902
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-41B15"
/clone_lib="RP11-11"
1..3439
/rpt_family="L1"
3444..3913
/rpt_family="ERV1"
4552..4773
/rpt_family="MER1_type"
5079..5100
/rpt_family="AT_rich"
5248..5393
/rpt_family="L2"
5433..5493
/rpt_family="L2"
6878..6898
/rpt_family="AT_rich"
6918..7076
/rpt_family="L1"
7077..7376
/rpt_family="MER2_type"
7377..7421
/rpt_family="L1"
7424..7956
/rpt_family="L1"
7950..9527
/rpt_family="L1"
9535..9659
/rpt_family="MALR"
9699..9901
/rpt_family="MER1_type"
9902..10204
/rpt_family="MALR"
10216..10241
/rpt_family="L1"
10246..10584
/rpt_family="L1"
10706..11007
/rpt_family="L1"

repeat_region 10989..11236
/rpt_family="L1"
repeat_region 11313..12046
/rpt_family="L1"
repeat_region 12141..12800
/rpt_family="L1"
repeat_region 13250..13342
/rpt_family="MIR"
repeat_region 13361..13391
/rpt_family="L1"
repeat_region 13392..14775
/rpt_family="L1"
repeat_region 14776..14922
/rpt_family="L1"
repeat_region 15274..15366
/rpt_family="L1"
repeat_region 15504..15670
/rpt_family="L1"
repeat_region 16039..16275
/rpt_family="L2"
repeat_region 16651..16785
/rpt_family="L1"
repeat_region 16855..17461
/rpt_family="L1"
repeat_region 17533..18613
/rpt_family="L1"
repeat_region 18738..19039
/rpt_family="L2"
repeat_region 19188..19659
/rpt_family="MALR"
repeat_region 19727..19815
/rpt_family="ERV1"
repeat_region 19890..20160
/rpt_family="L1"
repeat_region 20161..20461
/rpt_family="ALU"
repeat_region 20462..21332
/rpt_family="L1"
repeat_region 21337..21658
/rpt_family="L1"
repeat_region 21667..21728
/rpt_family="L1"
repeat_region 21729..22024
/rpt_family="ALU"
repeat_region 22025..22684
/rpt_family="L1"
repeat_region 22704..22887
/rpt_family="L1"
repeat_region 22888..23174
/rpt_family="ALU"
repeat_region 23175..24161
/rpt_family="L1"
24825..24828
/note="similar to Homo sapiens EST AA173493 (MID:91753661)"
zpo2904.t1
25016..25339
/rpt_family="L2"
25340..25520
/rpt_family="MIR"
25784..25812
/rpt_family="TCTCCC)n"
26401..26501
/rpt_family="L2"
26546..26629
/rpt_family="GA-rich"
27069..27173
/rpt_family="MIR"
28701..28784
/rpt_family="AT_rich"
29258..29502
/rpt_family="MALR"
29581..30086

REFERENCE 3 (bases 1 to 65335)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 65335)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 65335)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 6 (bases 1 to 65335)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jun 5, 2002 this sequence version replaced g1:20340530.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics -----
Center project name: H_NH0692E14

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-451F20, 2000 bp overlap;
the clone sequenced to the right is RP11-367N14. Actual start of
this clone is at base position 114362 of RP11-451F20.

There is a single plasmid region from 24260 to 24349.
Polymorphisms have been identified between AC107214 and AC112698.
Data from AC114911 was used to finish this clone, AC112698.
Location/Qualifiers
1. .65335
/organism="Homo sapiens"

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-692B14"
/clone_lib="RPCI-11"
410. .571
/rpt_family="MIR"
1465. .1792
/rpt_family="Alu"
2272. .2563
/rpt_family="MafR"
3341. .3644
/rpt_family="Alu"
3932. .4194
/rpt_family="MER2_type"
4136. .4356
/rpt_family="Alu"
4357. .4676
/rpt_family="Alu"
4689. .4739
/rpt_family="ERV1"
4757. .5034
/rpt_family="Alu"
5035. .5309
/rpt_family="L1"
5897. .6262
/rpt_family="CT-rich"
6264. .6334
/rpt_family="T-rich"
6824. .6937
/rpt_family="MIR"
7324. .7368
/rpt_family="(CA)n"
7861. .7893
/rpt_family="(TG)n"
8695. .8781
/rpt_family="MIR"
9092. .9128
/rpt_family="ERV1"
9129. .9435
/rpt_family="Alu"
9436. .9480
/rpt_family="ERV1"
9458. .9789
/rpt_family="ERV1"
10110. .10145
/rpt_family="L2"
10148. .10445
/rpt_family="Alu"
10793. .11080
/rpt_family="Alu"
12115. .12543
/rpt_family="L2"
12586. .12709
/rpt_family="Alu"
12720. .13003
/rpt_family="Alu"
13005. .13018
/rpt_family="(CA)n"
13019. .13053
/rpt_family="(CA)n"
13091. .13212
/rpt_family="Alu"
13245. .14096
/rpt_family="L2"
14285. .14596
/rpt_family="Alu"
16044. .16336
/rpt_family="Alu"
16463. .16779
/rpt_family="Alu"
16828. .16887
repeat_region

```
/rpt_family="MALR"
16892. .16976
/rpt_family="Alu"
17865. .17961
/rpt_family="CT-rich"
18410. .18722
/rpt_family="Alu"
18947. .19235
/rpt_family="Alu"
19442. .19590
/rpt_family="MER1_type"
19936. .20233
/rpt_family="Alu"
20486. .20628
/rpt_family="MIR"
21697. .21999
/rpt_family="Alu"
22113. .22196
/rpt_family="ERV1"
22197. .22481
/rpt_family="Alu"
22482. .22956
/rpt_family="ERV1"
23014. .23315
/rpt_family="Alu"
23397. .23523
/rpt_family="L2"
23623. .23947
/rpt_family="MER2_type"
24060. .24838
/rpt_family="L1"
24840. .25268
/rpt_family="MALR"
25269. .25457
/rpt_family="L1"
```

Query Match 19.5%: Score 202.2; DB 9; Length 6535;
Best Local Similarity 78.1%; Pred. No. 1.1e-29;
Matches 243; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```
QY 723 AGATACCTTAAGTTTATGTCACAGGCGTGGCGCATAGCTGATGCTGTGTCCTCCAGTGTCT 782
Db 50458 AGAATTTGAGTCTGAGAACAGAGGCTGGTGTGCTACGCTGTATCCCAACACT 50399
QY 783 TTGCGGGCCGACGACAGTGTGATCACTTAAGTCAAGATTGAGACCAAGCTGCCCCAAC 842
Db 50398 TTGGAGGCTGAGGACAGGTGATCACCTGAGTCAAGAGTTCGAGACCAAGTCTGSCCTAG 50339
QY 843 ATGGTGAACGTTGCTCTACTAATAAAATACAAAATTAGACAGGCGTGTGGCAACATC 902
Db 50338 ATGGTGAACCCCACTCTCTACTAATAAAATACAAAATTAGCTGGGTGTGATGTGGCACCC 50279
QY 903 TGTAAATTCAGCTACTCAGAGGCTAACACAGAAATTCCTGAACCTGGAGGACAGAG 962
Db 50278 TGTAAATCCAGCTCTTAAGAGGCGCAAGGACGAATCGCTGAACCGGAGGACAGAT 50219
QY 963 GTTGCAGTAGCCATTGCACTTCAGCTGGGCAACACAGTAGACTCTTGTCTCAAAAAA 1022
Db 50218 GTTGCAGTAGACCCGCTGCACTCCAGCTGGGCAACATTAACAGATCTGTCTCAAAAAA 50159
QY 1023 AAAAAAAAAA 1033
Db 50158 AAAAAAAAAA 50148
```

Search completed: March 6, 2005, 05:42:31
Job time : 4571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 04:05:58 ; Search time 600 Seconds
(Without alignments)
10241.153 Million cell updates/sec

Title: US-10-664-358-36

Perfect score: 1038

Sequence: 1 gnaatcgcgcagcagtcac.....aaaaaaaaaactcga 1038

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980b:*
2: geneseqn1980b:*
3: geneseqn2000b:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1035	99.7	1038	2	AAX51747 DNA encod
2	1035	99.7	1038	6	ABO92599 Human sec
3	1035	99.7	1038	8	ACC50369 Human sec
4	1035	99.7	1038	8	ABZ71204 Human sec
5	1035	99.7	1038	9	ADB91088 Human sec
6	1035	99.7	1038	10	ADC73416 Human sec
7	200	19.3	68196	13	ABD32800 Human can
8	199.8	19.2	122748	6	ABT10719 Human bre
9	199.2	19.2	92500	12	ADP45591 Human int
10	199.2	19.2	110000	10	ACF42745 Human ALM
11	199.2	19.2	310268	13	ABD32548 Human can
12	198.2	19.1	32249	4	AA104931 Human tes
13	198.2	19.1	32249	4	ABL97825 Human rep
14	196.6	18.9	4165	9	AA161140 Human 76P
15	196.6	18.9	4165	12	ADM98105 Human nat
16	196.6	18.9	6837	9	AA161118 Human 76P
17	196.6	18.9	6837	12	ADM98083 Human nat
18	196.6	18.9	10394	12	AA161117 Human tum
19	196.6	18.9	10394	12	ADM98082 Human tum
20	196.4	18.9	4297	4	AA541953 Genomic B

C	21	196.4	18.9	6072	10	ADFe9160	Adfe9160 Human MP5
C	22	196.2	18.9	122186	4	AAc99560	AAc99560 Human his
C	23	196	18.9	6239	4	AAK76233	AAK76233 Human imm
C	24	196	18.9	167343	6	ABLe4403	ABLe4403 Stomach c
C	25	196	18.9	167343	6	ABLe7239	ABLe7239 Thyroid c
C	26	195.6	18.8	65464	8	ABX13172	ABX13172 Human gen
C	27	195.4	18.8	18981	4	AA104961	AA104961 Human rep
C	28	195.4	18.8	18981	4	AB197854	AB197854 Human tes
C	29	195.4	18.8	109586	11	ACN43994	ACN43994 Human gen
C	30	195	18.8	1160	4	AA102996	AA102996 Human rep
C	31	195	18.8	1160	4	AB197343	AB197343 Human tes
C	32	195	18.8	254396	12	ADQ97557	ADQ97557 Human can
C	33	194.8	18.8	3939	8	AA150288	AA150288 Human nox
C	34	194.4	18.7	68255	11	ACN44362	ACN44362 Human gen
C	35	194.4	18.7	110000	11	ACN44582_0	ACN44582 Human gen
C	36	194.4	18.7	110000	13	ABD32535_0	ABD32535 Human can
C	37	193.8	18.7	22062	12	ADQ97922	ADQ97922 Human can
C	38	193.8	18.7	110000	11	ACN43998_5	ACN43998 (6 of
C	39	193.6	18.7	72705	11	ACN45158	ACN45158 Human gen
C	40	193.6	18.7	104062	11	ACN44458	ACN44458 Human gen
C	41	193.4	18.6	220756	12	ADG86300	ADG86300 Human SKR
C	42	193.4	18.6	233380	11	ACN44282	ACN44282 Human gen
C	43	193.2	18.6	110000	11	ACN43998_4	ACN43998 (5 of
C	44	193	18.6	6644	5	AAF31479	AAF31479 ZALPHA as
C	45	193	18.6	6644	10	ADP18591	ADP18591 Human zal

ALIGNMENTS

RESULT 1
AAX51747
ID AAX51747 standard; DNA, 1038 BP.
XX
AC AAX51747;
DT 17-JUN-1999 (first entry)
XX
DB DNA encoding a human secreted protein.
XX
XX Human secreted protein; cancer; immune disorder; infection;
KW inflammatory disorder; skin disorder; tumour; atherosclerosis;
KW restenosis; autoimmune disorder; Alzheimer's disease;
KW peripheral neuropathy; trauma; spinal cord injury; allergy;
KW hematopoietic disorder; skeletal disorder; neurological disorder;
KW arthritic disorder; asthma; immunodeficiency disease; AIDS;
transplant rejection; ss.
XX
OS Homo sapiens.
XX
XX W09911293-AL.
XX
XX 11-MAR-1999.
XX
XX 03-SEP-1998; 98WC-US018360.
XX
XX 05-SEP-1997; 97US-0057626P.
XX
XX 05-SEP-1997; 97US-0057663P.
XX
XX 05-SEP-1997; 97US-0057663P.
XX
XX 12-SEP-1997; 97US-0058666P.
XX
XX 12-SEP-1997; 97US-0058667P.
XX
XX 12-SEP-1997; 97US-0058973P.
XX
XX 12-SEP-1997; 97US-0058974P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen GA, Olsen HS;
XX Ebner R, Brewer LA;
XX WPI; 1999-204988/17.
XX P-PSDB; AAY12960.
XX
XX New isolated human genes and the secreted polypeptides they encode -

PT useful for diagnosis and treatment of e.g. neurological disorders,
PT tumors, immune disorders, inflammation or hematological disorders.
XX
PS Claim 1, Page 176-177, 215pp; English.
XX
CC AAX51701-55 encode human secreted proteins. The polynucleotides and their
CC corresponding secreted polypeptides are useful for preventing, treating
CC or ameliorating medical conditions, e.g. by protein or gene therapy.
CC Pathological conditions can also be diagnosed by determining the amount
CC of the new polypeptides in a sample or by determining the presence of
CC mutations in the new polynucleotides. Specific uses are described for
CC each polynucleotide, based on which tissues they are most highly
CC expressed in, and include developing products for the diagnosis or
CC treatment of cancer, immune disorders, infection, inflammatory disorders,
CC skin disorders, tumours, atherosclerosis, restenosis, autoimmune
CC disorders, Alzheimer's disease, peripheral neuropathies, trauma, spinal
CC cord injuries, allergy, hematopoietic disorders, skeletal disorders,
CC neurological disorders, arthritic disorders, asthma, immunodeficiency
CC diseases, AIDS and transplant rejection. The polypeptides are also useful
CC for identifying their binding partners
XX
SQ Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;

Query Match 99.7%; Score 1035; DB 2; Length 1038;
Best Local Similarity 100.0%; Pred. No. 4.2e-213;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GNAATTGGCAGAGTTAATGTAATAATTTCTATATAATTTTAAAGGAATTGA 60
DB 1 GNAATTGGCAGAGTTAATGTAATAATTTCTATATAATTTTAAAGGAATTGA 60
QY 61 GCATCATGAAAAAATGCTCTTACTGTTGAAAAACATATTTGTATACATTTTGGTCAACTA 120
DB 61 GCATCATGAAAAAATGCTCTTACTGTTGAAAAACATATTTGTATACATTTTGGTCAACTA 120
QY 121 ATCTTCAATTAATCTTTTGTAACTATATTAATGTTAGTTCACAGTGGCAGTCTTATATAG 180
DB 121 ATCTTCAATTAATCTTTTGTAACTATATTAATGTTAGTTCACAGTGGCAGTCTTATATAG 180
QY 181 TAAATGGCAGCTGACAGATGAAAAATACATATCTTAATTTTGTACTATCTTATAGG 240
DB 181 TAAATGGCAGCTGACAGATGAAAAATACATATCTTAATTTTGTACTATCTTATAGG 240
QY 241 AAAATCAGAGATTTCAAAACCTTGTATGTTTAAAGGTATAGTACATTTTATATATGT 300
DB 241 AAAATCAGAGATTTCAAAACCTTGTATGTTTAAAGGTATAGTACATTTTATATATGT 300
QY 301 GCGGTATTTTATACATGATTTGACGTTTGMAAATATTTTCCCTGGACTTTTATTTTA 360
DB 301 GCGGTATTTTATACATGATTTGACGTTTGMAAATATTTTCCCTGGACTTTTATTTTA 360
QY 361 GATGAGATCTACAGTGTAGGCAAACTTATATATCTGTCAACATTCATTAGTGCATAGTC 420
DB 361 GATGAGATCTACAGTGTAGGCAAACTTATATATCTGTCAACATTCATTAGTGCATAGTC 420
QY 421 AGACTCATCCCATGCTTAAATTAATTAAGTTGKAAATACGTTTGTAAATAGTTGTGT 480
DB 421 AGACTCATCCCATGCTTAAATTAATTAAGTTGKAAATACGTTTGTAAATAGTTGTGT 480
QY 481 AGGTCAATTAACCAAGCTTCAAGKATTAATTAATTAATTAATTAATTAATTAATTAAT 540
DB 481 AGGTCAATTAACCAAGCTTCAAGKATTAATTAATTAATTAATTAATTAATTAATTAAT 540
QY 541 TGAATAMCCGTTTTTCCATGCAAAAGTAAATCTTCAGCCTTTAATTTTTTATTAAT 600
DB 541 TGAATAMCCGTTTTTCCATGCAAAAGTAAATCTTCAGCCTTTAATTTTTTATTAAT 600
QY 601 ATATTAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 ATATTAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GCTTTGCTAAATTTGTAACCTGACAAATCTTACGAGTCTTCATTTTGTGAGAT 720
DB 661 GCTTTGCTAAATTTGTAACCTGACAAATCTTACGAGTCTTCATTTTGTGAGAT 720
```

```
DB 661 GCTTTGCTAAATTTGTAACCTGACAAATCTTACGAGTCTTCATTTTGTGAGAT 720
QY 721 GAAGATACCTTAGTTTATGTCAGAGGCTGGGCGGATAGCTGATGCTGTGTCAGAGT 780
DB 721 GAAGATACCTTAGTTTATGTCAGAGGCTGGGCGGATAGCTGATGCTGTGTCAGAGT 780
QY 781 CTTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
DB 781 CTTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 841 ACATGTGAACGTTGCTCTACTATTAATAATAATAATAATAATAATAATAATAATAATA 900
DB 841 ACATGTGAACGTTGCTCTACTATTAATAATAATAATAATAATAATAATAATAATAATA 900
QY 901 TCTGTAAATTCAGTCTACTACAGAGGCTTAAACAGAGAAATTCCTTGAACCTGGAGGCG 960
DB 901 TCTGTAAATTCAGTCTACTACAGAGGCTTAAACAGAGAAATTCCTTGAACCTGGAGGCG 960
QY 961 AGGTGACAGGACGCTTGAACCTGACCTGCGGCAACAGTGAAGCTTGTCTCAAAA 1020
DB 961 AGGTGACAGGACGCTTGAACCTGACCTGCGGCAACAGTGAAGCTTGTCTCAAAA 1020
QY 1021 AAAAAAAAAAACTCGA 1038
DB 1021 AAAAAAAAAAACTCGA 1038
```

RESULT 2
AB092599
ID AB092599 standard; cDNA; 1038 BP.
XX
AC AB092599;
XX
DT 12-NOV-2002 (first entry)
XX
DE Human secreted protein encoding cDNA SEQ ID NO 57.
XX
KW Human; nootropic; neuroprotective; cytosolic; dermatological; virucide;
KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antiskilling; antianaemic; antirheumatic; cancer;
KW antihemetic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antifungal; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200257420-A2.
XX
PD 25-JUL-2002.
XX
PF 17-JAN-2002; 2002MO-US001109.
XX
PR 18-JAN-2001; 2001US-0262066P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen CA, Olsen H,
PI Ebner R, Brewer LA;
XX
DR P-PSDB; ABP62059.
XX
WI PI; 2002-599716/64.
XX
PT New polynucleotides and polypeptides useful for diagnosing, prognosing,
PT treating or preventing e.g. neurodegenerative, central nervous system,
PT autoimmune, respiratory, reproductive, or inflammatory diseases or
PT disorders.
XX
PS Claim 1, Page 726; 785pp; English.
XX
CC The invention relates to novel genes (AB092553-AB092607) and proteins
CC (ABP62013-ABP62153) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections
CC
XX

80 Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;

Query Match 99.7%; Score 1035; DB 6; Length 1038;
Best Local Similarity 100.0%; Pred. No. 4.2e-213;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GNAATGGGCGACGAGTTAATGATATAATTTCTATATGAATTTAAATGGGAATTGA 60
Db 1 GNAATGGGCGACGAGTTAATGATATAATTTCTATATGAATTTAAATGGGAATTGA 60
Qy 61 GCATCATAGAAAAAATGCTCTTACTGTGAAAAACATTATTTGTACATTTTGTCAACTA 120
Db 61 GCATCATAGAAAAAATGCTCTTACTGTGAAAAACATTATTTGTACATTTTGTCAACTA 120
Qy 121 ATCTTCAATACCTTTTAGTAACTATATGTTAAGTTGACAGTGGCAGCTTATATAG 180
Db 121 ATCTTCAATACCTTTTAGTAACTATATGTTAAGTTGACAGTGGCAGCTTATATAG 180
Qy 181 TAAATGGCAGCTGACAGCATGAAATAACATCTATATTTTGTGCTATCTTATAGG 240
Db 181 TAAATGGCAGCTGACAGCATGAAATAACATCTATATTTTGTGCTATCTTATAGG 240
Qy 241 AAAATCAGAGATTTCAAAAACCTTGTAGTTTATAGGTAATGACATTTTATATAATGT 300
Db 241 AAAATCAGAGATTTCAAAAACCTTGTAGTTTATAGGTAATGACATTTTATATAATGT 300
Qy 301 GCGGTAATTTATATCATGATTTGACGTTTGTMAAATATTTTCCGTGACATTTATTTTA 360
Db 301 GCGGTAATTTATATCATGATTTGACGTTTGTMAAATATTTTCCGTGACATTTATTTTA 360
Qy 361 GATGAGATCTACAGTGTAGGCAAACTTATATATCTGTCACTCCATTAAGTGTATGTC 420
Db 361 GATGAGATCTACAGTGTAGGCAAACTTATATATCTGTCACTCCATTAAGTGTATGTC 420
Qy 421 AGACTCATCCCATGCTAAATATTAAGTTGTAAATACGCTTTGTAAATAGTTGTGTT 480
Db 421 AGACTCATCCCATGCTAAATATTAAGTTGTAAATACGCTTTGTAAATAGTTGTGTT 480
Qy 481 AGGTCATTAACCAAGCTTTCAAGKATTAATTATATAAAAACCTTGKTTTATCTTG 540
Db 481 AGGTCATTAACCAAGCTTTCAAGKATTAATTATATAAAAACCTTGKTTTATCTTG 540
Qy 541 TGAATATCCGTTTTTCCATGCAAAAGTTAAATCTTCAGCTTTAATTTTATATAT 600
Db 541 TGAATATCCGTTTTTCCATGCAAAAGTTAAATCTTCAGCTTTAATTTTATATAT 600
Qy 601 ATATTAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 601 ATATTAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 661 GCTTTTCTAATATGTTACCTGACAAAATCTTACGAGTTCTTCAATTTTGGTTGAGAT 720
Db 661 GCTTTTCTAATATGTTACCTGACAAAATCTTACGAGTTCTTCAATTTTGGTTGAGAT 720
Qy 721 GAAGATATCTAGTTTGTGTCAGAGGGCTGGCGGAGATAGCTGAGCTGCTGGTCCAGTG 780
Db 721 GAAGATATCTAGTTTGTGTCAGAGGGCTGGCGGAGATAGCTGAGCTGCTGGTCCAGTG 780
Qy 781 CTTTGCGGGCCGAGGAGGTGATCACTTAAGGTCAGAGGTTGAGACACAGCTGCGCA 840

Db 781 CTTTGCGGGCCGAGGAGGTGATCACTTAAGGTCAGAGGTTGAGACACAGCTGCGCA 840
Qy 841 ACATGTGAAACCTGTCTTACTATAAAATATCAAAATTTAGACAGCGGTGTGCAACA 900
Db 841 ACATGTGAAACCTGTCTTACTATAAAATATCAAAATTTAGACAGCGGTGTGCAACA 900
Qy 901 TCTGTAATTCAGCTACTCAGAGAGCTTACACAGAAAAATCTTGAACCTGGAGGAG 960
Db 901 TCTGTAATTCAGCTACTCAGAGAGCTTACACAGAAAAATCTTGAACCTGGAGGAG 960
Qy 961 AGGTTGAGTGAGGCACTTGGCACTCGAGGCAACAGTGAAGACTTTGTCTCAAAA 1020
Db 961 AGGTTGAGTGAGGCACTTGGCACTCGAGGCAACAGTGAAGACTTTGTCTCAAAA 1020
Qy 1021 AAAAAAAAAAAACTCGA 1038
Db 1021 AAAAAAAAAAAACTCGA 1038

RESULT 3

ID ACC50369 standard; cDNA; 1038 BP.
XX
XX ACC50369;
XX
XX 12-JUN-2003 (first entry)
XX
XX Human secreted protein coding sequence, SEQ ID 36.
XX
XX Cardiac; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;
XX vulnerable; antiinflammatory; neurotropic; neuroprotective;
XX antiparkinsonian; gene therapy; human; cardiovascular disorder; gene; ss.
XX Homo sapiens.
XX
XX WO200295010-A2.
XX
XX 28-NOV-2002.
XX
XX 19-MAR-2002; 2002WO-US009785.
XX
XX 21-MAR-2001; 2001US-0277340P.
XX 19-JUL-2001; 2001US-0306171P.
XX 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-129429/12.
XX
XX Novel human secreted proteins, useful for detecting, preventing,
XX diagnosing, prognosticating, treating and/or ameliorating cardiovascular
XX disorders such as arrhythmia.
XX
XX Claim 21; SEQ ID NO 36; 1881bp; English.
XX
XX The present invention relates to novel human secreted proteins (ABR4/633-
XX ABR4/8145) and their coding sequences (ACC50344-ACC50856). The proteins
XX and their coding sequences are useful for the preparation of a diagnostic
XX or pharmaceutical composition for diagnosing or treating a cardiovascular
XX disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
XX arteriosclerosis and myocardial ischaemia), neural disorders, immune
XX system disorders, muscular disorders, reproductive disorders,
XX gastrointestinal disorders, pulmonary disorders, renal disorders,
XX proliferative disorders and/or cancerous diseases and conditions, for
XX wound healing and epithelial cell proliferation, to treat inflammation or
XX infection, for treating thrombosis and arteriosclerosis, for treating or
XX preventing neural damage which occurs in neuronal disorders or
XX neurodegenerative conditions such as Alzheimer's disease and Parkinson's
XX disease, to enhance bone and periodontal regeneration and aid in tissue
XX transplants or bone grafts, to prevent skin aging or hair loss, to

CC stimulate growth and differentiation of haematopoietic cells and bone
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism. Note: The sequence data for this patent was published in
CC electronic format and is available from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;

Query Match 99.7%; Score 1035; DB 8; Length 1038;
Best Local Similarity 100.0%; Pred. No. 4.2e-213;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GNAATTCGGACGACGATTAATGTAATAAATTTCTTAATGAATTTTAATGGCAATTAGA 60
DB 1 GNAATTCGGACGACGATTAATGTAATAAATTTCTTAATGAATTTTAATGGCAATTAGA 60
QY 61 GCATCATAGAAAAAATGCTCTTACTGTGAAAAAATTATTTGTTACATTTTGGTCAACTA 120
DB 61 GCATCATAGAAAAAATGCTCTTACTGTGAAAAAATTATTTGTTACATTTTGGTCAACTA 120
QY 121 ATCTTTCATTAATCTTTTACTTAATATGTTAAGTTGACAGTGGCAGTCTTATATAG 180
DB 121 ATCTTTCATTAATCTTTTACTTAATATGTTAAGTTGACAGTGGCAGTCTTATATAG 180
QY 181 TAAATGGACGTGACGATGAAATTAATCTTAATTTTGTGACATCTTATTTAG 240
DB 181 TAAATGGACGTGACGATGAAATTAATCTTAATTTTGTGACATCTTATTTAG 240
QY 241 AAAATCAGAAATTTCAAAACCTTGTAGTTTAAAGGATAGTACATTTTAAATAGT 300
DB 241 AAAATCAGAAATTTCAAAACCTTGTAGTTTAAAGGATAGTACATTTTAAATAGT 300
QY 301 GCGGTATATTTATACATGATTTGAGTTGTGMAAATTTTCCCTGGAATTTATTTTA 360
DB 301 GCGGTATATTTATACATGATTTGAGTTGTGMAAATTTTCCCTGGAATTTATTTTA 360
QY 361 GATGAGATCTACAGTGTAGGCAAACTTAATATCTGCACTCATTAGTGTATGTC 420
DB 361 GATGAGATCTACAGTGTAGGCAAACTTAATATCTGCACTCATTAGTGTATGTC 420
QY 421 AGACTCATCCCGATGCTAAATTAATATGTTGKAAATAACGTTTGTAAATAGTTGTT 480
DB 421 AGACTCATCCCGATGCTAAATTAATATGTTGKAAATAACGTTTGTAAATAGTTGTT 480
QY 481 AGGTCAATATCAACAAGCTTCAAGGKATTAATTAATAAACCCTGKTTTATTTCTTG 540
DB 481 AGGTCAATATCAACAAGCTTCAAGGKATTAATTAATAAACCCTGKTTTATTTCTTG 540
QY 541 TGAATAMCCGTTTTTTCAGTCAAAAGTTAAATCTTCAGCCTTTAATTTTAAATTAAT 600
DB 541 TGAATAMCCGTTTTTTCAGTCAAAAGTTAAATCTTCAGCCTTTAATTTTAAATTAAT 600
QY 601 ATATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 ATATAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 GCTTTTGTAAATTTTACTGACGAAATCTTAGCCAGTCTTCAATTTTGTTTGAAT 720
DB 661 GCTTTTGTAAATTTTACTGACGAAATCTTAGCCAGTCTTCAATTTTGTTTGAAT 720
QY 721 GAAATATCTTAATTTTGTGTCAGGGGCTGGGCGGATAGCTGATGCTGTGTCAGATG 780
DB 721 GAAATATCTTAATTTTGTGTCAGGGGCTGGGCGGATAGCTGATGCTGTGTCAGATG 780
QY 781 CTTTGCGGGGCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 781 CTTTGCGGGGCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 841 ACAATGTAAGCTTGTCTTACTTAATAAATATGACAGGCGTGTGTCACACA 900
DB 841 ACAATGTAAGCTTGTCTTACTTAATAAATATGACAGGCGTGTGTCACACA 900
```

```
DB 841 ACATGTTGAAAAGTTGCTCTTACTTAATAAATATGACAGGCGTGTGTCACACA 900
QY 901 TCTGTAATTTCCAGCTACTCAGAGGCTTAACAACAGAAAATTTCTTGAATCTGGAGGACAG 960
DB 901 TCTGTAATTTCCAGCTACTCAGAGGCTTAACAACAGAAAATTTCTTGAATCTGGAGGACAG 960
QY 961 AGTTGACGTAGGCACTTGTGACCTTCAGCCTTGGGCAACAGATGAGACTTGTCTCAAAA 1020
DB 961 AGTTGACGTAGGCACTTGTGACCTTCAGCCTTGGGCAACAGATGAGACTTGTCTCAAAA 1020
QY 1021 AAAAAAAAAAAACTGCA 1038
DB 1021 AAAAAAAAAAAACTGCA 1038

RESULT 4
ABZ71204
ID ABZ71204 standard; cDNA; 1038 BP.
XX
AC ABZ71204;
XX
DT 03-Apr-2003 (first entry)
XX
XX Human secreted protein-encoding gene 15 cDNA clone HB1AE26, SEQ ID NO:25.
KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic;
KW antiinflammatory; immunosuppressive; vulnerability; gene therapy; gene; ss.
OS Homo sapiens.
XX
PN NC0200276488-A1.
XX
PD 03-OCT-2002.
XX
PE 19-MAR-2002; 2002MO-US008276.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-029900/02.
XX
XX P-PSDB; ABR00025.
XX
XX New human secreted proteins and nucleic acids, useful for detecting,
XX preventing, diagnosing, prognosticating, treating and/or ameliorating
XX e.g. gastrointestinal diseases and disorders, or cancers.
XX
XX Claim 21; Page 764-765; 1216pp; English.
XX
XX ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
XX protein genes, and ABP00011-ABP00299 represent the proteins they encode.
XX ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
XX invention also encompasses antibodies specific for the secreted proteins,
XX the use of the secreted proteins in drug screening, and recombinant
XX vectors and host cells comprising a nucleic acid of the invention. The
XX secreted proteins, nucleic acids encoding them, antibodies or antibody
XX fragments are useful for the secreted proteins, treating, and modulators of protein
XX digestive disorders. Such conditions include disorders of the mouth,
XX oesophagus, stomach, small intestine, large intestine, liver, biliary
XX tract and pancreas, and include cancers of these organs and tissues. The
XX secreted proteins and their nucleic acids may also be used in the
XX treatment of immune disorders, inflammation, infection,
XX hyperproliferative disorders, and to promote wound healing. Nucleic acids
XX of the invention may be used for chromosome identification, chromosome
```

CC mapping, in gene therapy, for identifying individuals from minute
CC biological samples, as hybridisation probes, and as molecular weight
CC markers. The present sequence represents a human secreted protein-
CC encoding cDNA clone of the invention

XX Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;

Query Match 99.7%; Score 1035; DB 8; Length 1038;
Best Local Similarity 100.0%; Pred. No. 4.2e-213;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GNAATTCGGCAGCAGTTAATGATATAAATTTCTATATGAATTTTAAATGGAATTAGA 60
DB 1 GNAATTCGGCAGCAGTTAATGATATAAATTTCTATATGAATTTTAAATGGAATTAGA 60
QY 61 GCATCATAGAAAAAATGCTCTTACTGTGGAAAAATATTGTTACATTTGGTCAACTA 120
DB 61 GCATCATAGAAAAAATGCTCTTACTGTGGAAAAATATTGTTACATTTGGTCAACTA 120
QY 121 ATCTTTCAATTAATCTTTAGTAACTATATAATGTTAAGTTGACAGTGGCAGTCTTATATAG 180
DB 121 ATCTTTCAATTAATCTTTAGTAACTATATAATGTTAAGTTGACAGTGGCAGTCTTATATAG 180
QY 181 TAAATGGCAGCTGACAGCATGAAAAATACATATCTAATATTTTGTACTATCTTATAGG 240
DB 181 TAAATGGCAGCTGACAGCATGAAAAATACATATCTAATATTTTGTACTATCTTATAGG 240
QY 241 AAAATCAGAGAAATTCAAAACCTTGTTAGTTTAAAGGTATAGTCACTTTTAAATATGT 300
DB 241 AAAATCAGAGAAATTCAAAACCTTGTTAGTTTAAAGGTATAGTCACTTTTAAATATGT 300
QY 301 GCGGTATATTTATACATGATTTGACGTTTGTGMAAATTTTCCCTGGACCTTTATTTTA 360
DB 301 GCGGTATATTTATACATGATTTGACGTTTGTGMAAATTTTCCCTGGACCTTTATTTTA 360
QY 361 GATGAGATCTACAGTGTAGGCAAACTTATATATCTGTCAACTGCATTAGTGTATAGTC 420
DB 361 GATGAGATCTACAGTGTAGGCAAACTTATATATCTGTCAACTGCATTAGTGTATAGTC 420
QY 421 AGACTCATCCCGCAGCTTAAATTTATAGTTGTAAAAATACGTTTGTGTAATAGTTGTG 480
DB 421 AGACTCATCCCGCAGCTTAAATTTATAGTTGTAAAAATACGTTTGTGTAATAGTTGTG 480
QY 481 AGGTCATTATCACAAGCTTCAAGGKATATCATATTAATAAAGCTTGKTTTATTCGTG 540
DB 481 AGGTCATTATCACAAGCTTCAAGGKATATCATATTAATAAAGCTTGKTTTATTCGTG 540
QY 541 TGAATAMCCGTTTTTTCATGCAAAAGTTAAATCTTCAGCCTTTAAATTTTATTATAT 600
DB 541 TGAATAMCCGTTTTTTCATGCAAAAGTTAAATCTTCAGCCTTTAAATTTTATTATAT 600
QY 601 ATATTAAGATGTGATGATAGTACTACACAAAACAGAAAAAATAACAGATTCGTTGTG 660
DB 601 ATATTAAGATGTGATGATAGTACTACACAAAACAGAAAAAATAACAGATTCGTTGTG 660
QY 661 GCTTTTCTAAATTTGTTACCTGACAAAATCTTACCCAGTTTCTTCAATTTTCTTTTGAT 720
DB 661 GCTTTTCTAAATTTGTTACCTGACAAAATCTTACCCAGTTTCTTCAATTTTCTTTTGAT 720
QY 721 GAAGATACTAGTTTATAGTCCAGAGGGCTGGCGGATAGTGTGATGCTGTGCTCCAGTG 780
DB 721 GAAGATACTAGTTTATAGTCCAGAGGGCTGGCGGATAGTGTGATGCTGTGCTCCAGTG 780
QY 781 CTTTGGCGGGCGGAGGCGATGATCACTTAAGGTCAAGGATTTTGAAGCCAGCTGCCCA 840
DB 781 CTTTGGCGGGCGGAGGCGATGATCACTTAAGGTCAAGGATTTTGAAGCCAGCTGCCCA 840
QY 841 ACATGTGAAAACGTTGCTCTTACTATAAATAATCAAAATTTGAACGCGGTGTGGCACACA 900
DB 841 ACATGTGAAAACGTTGCTCTTACTATAAATAATCAAAATTTGAACGCGGTGTGGCACACA 900
QY 901 TCTGTATTTCCAGCTACTACAGAGAGCTTAACAGAAAAATTCCTTGAACTGTGGAGGACAG 960
DB 901 TCTGTATTTCCAGCTACTACAGAGAGCTTAACAGAAAAATTCCTTGAACTGTGGAGGACAG 960
```

```
DB 901 TCTGTATTTCCAGCTACTACAGAGAGCTTAACAGAAAAATTCCTTGAACTGTGGAGGACAG 960
QY 961 AGGTTGACGTAGAGCAATTTGCACTCCAGCTGGGCAACAGATGAGACTTGTCTCAAAA 1020
DB 961 AGGTTGACGTAGAGCAATTTGCACTCCAGCTGGGCAACAGATGAGACTTGTCTCAAAA 1020
QY 1021 AAAAAAAAAAAAACTCGA 1038
DB 1021 AAAAAAAAAAAAACTCGA 1038
```

RESULT 5

ADB91088
ID ADB91088 standard; cDNA; 1038 BP.

AC ADB91088;

DT 04-DEC-2003 (first entry)

DE Human secreted protein cDNA #SEQ ID 34.

KM Secreted protein; gene therapy; antidiabetic; diabetes; human; ss.

OS Homo sapiens.

PN WO200304622-A2.

PD 16-JAN-2003.

PE 19-MAR-2002; 2002MO-US008124.

PR 21-MAR-2001; 2001US-0277340P.

PR 19-JUL-2001; 2001US-0306171P.

PR 13-NOV-2001; 2001US-0331287P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

XX WPI; 2003-229407/22.

XX Nucleic acid encoding a human secreted protein is useful in diagnosing or
XX treating diabetes or conditions related to diabetes.

XX Claim 9; SEQ ID NO 34; 1537bp; English.

XX The invention relates to isolated nucleic acid molecules ADB91065-
XX ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
XX ADB91834. Also disclosed is a recombinant vector comprising a

XX polynucleotide of the invention, and a recombinant host cell comprising
XX the recombinant vector. The polypeptide of the invention is useful in

XX identifying a binding partner by contacting the polypeptide with a
XX binding partner, and determining whether the binding partner increases or

XX decreases activity of the polypeptide. The polypeptide, polynucleotide,
XX antibody or its fragment, agonist or antagonist are useful for preparing

XX a pharmaceutical composition for diagnosing or treating diabetes or
XX conditions related to diabetes. The present sequence is that of the human

XX immunoglobulin Fc portion used to generate fusion proteins, increasing
XX the stability of the fused protein as compared to the secreted protein

XX only. Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;

Query Match 99.7%; Score 1035; DB 9; Length 1038;
Best Local Similarity 100.0%; Pred. No. 4.2e-213;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GNAATTCGGCAGCAGTTAATGATATAAATTTCTATATGAATTTTAAATGGAATTAGA 60
DB 1 GNAATTCGGCAGCAGTTAATGATATAAATTTCTATATGAATTTTAAATGGAATTAGA 60
```

```
QY 61 GCATCATGAGAAAAATGCTCTTACTGTGAAAAACATATTTGTACATTTTGTCAACTA 120
DB 61 GCATCATGAGAAAAATGCTCTTACTGTGAAAAACATATTTGTACATTTTGTCAACTA 120
QY 121 ATCTTCAATTAACCTTTTAACTATATATGTTAAGTGTACAGATGCGAGCTTATATAG 180
DB 121 ATCTTCAATTAACCTTTTAACTATATATGTTAAGTGTGTACAGATGCGAGCTTATATAG 180
QY 181 TAAATGCGAGCTGACAGCATGAAAAATTAACATATCTATATTTTGTGACTATCTTATAGG 240
DB 181 TAAATGCGAGCTGACAGCATGAAAAATTAACATATCTATATTTTGTGACTATCTTATAGG 240
QY 241 AAAATCAGAGAAATTCAAAACCTTGTAGTTTAAAGGTAATGACATTTATATAATGT 300
DB 241 AAAATCAGAGAAATTCAAAACCTTGTAGTTTAAAGGTAATGACATTTATATAATGT 300
QY 301 GCGGTATATTTATACATGATTTGACGTTTGTGMAAATATTTCCCTGACATTTTATTTTA 360
DB 301 GCGGTATATTTATACATGATTTGACGTTTGTGMAAATATTTCCCTGACATTTTATTTTA 360
QY 361 GATGAGATCTACAGCTGAGCAAACTTATATATCTGCACTGCATTAAGTGTATAGTC 420
DB 361 GATGAGATCTACAGCTGAGCAAACTTATATATCTGCACTGCATTAAGTGTATAGTC 420
QY 421 AGACTCAATCCCATGCTAAAATATAGTGTGKAAAATACGCTTTTGTATAATGTTAGTT 480
DB 421 AGACTCAATCCCATGCTAAAATATAGTGTGKAAAATACGCTTTTGTATAATGTTAGTT 480
QY 481 AGGTCATTAATCACCAAGCTTCAAGGKATTAATTATATAAAACCTTGKTTTATCTTG 540
DB 481 AGGTCATTAATCACCAAGCTTCAAGGKATTAATTATATAAAACCTTGKTTTATCTTG 540
QY 541 TGAATATCCGTTTTTTCATGCAAAAGTTAAATCTTCAAGCTTTATTTTATTTATAT 600
DB 541 TGAATATCCGTTTTTTCATGCAAAAGTTAAATCTTCAAGCTTTATTTTATTTATAT 600
QY 601 ATATTAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 ATATTAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GCTTTTGTCTAAATTTGTTACCTGACAAAATCTTACGAGTTCTTCAATTTTGTGAGAT 720
DB 661 GCTTTTGTCTAAATTTGTTACCTGACAAAATCTTACGAGTTCTTCAATTTTGTGAGAT 720
QY 721 GAAGATCTTAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 721 GAAGATCTTAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 CTTTGGCGGGGCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 781 CTTTGGCGGGGCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 841 ACATGATGAAAAGTTGTCTCTACTTAAATAAATAAATAATTAAGACAGGCGTGTGCACACA 900
DB 841 ACATGATGAAAAGTTGTCTCTACTTAAATAAATAAATAATTAAGACAGGCGTGTGCACACA 900
QY 901 TCTGTATTTCAAGTACTACAGAGAGCTTAAAGTCAAGAGTTGAAGCCAGCTGCCCA 960
DB 901 TCTGTATTTCAAGTACTACAGAGAGCTTAAAGTCAAGAGTTGAAGCCAGCTGCCCA 960
QY 961 AGGTTGAGAGTGAAGCATTTGACCTCAGCTGCGGCAACAGATGAGACTTGTCTCAAAA 1020
DB 961 AGGTTGAGAGTGAAGCATTTGACCTCAGCTGCGGCAACAGATGAGACTTGTCTCAAAA 1020
QY 1021 AAAAAAAAAAAAACTCGA 1038
DB 1021 AAAAAAAAAAAAACTCGA 1038
RESULT 6
ADC73416
ID ADC73416 standard; DNA; 1038 BP.
XX
```

```
AC ADC73416;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human secreted protein-related DNA - SEQ ID 49.
XX
KW antidiabetic; antineuronic; antiarthritic; antiinflammatory; antihypertoid;
KW antidiabetic; immunosuppressive; dermatologic; nephrotoxic;
KW antiparkinsonian; neuroprotective; nootropic; antibacterial; virocidic;
KW fungicide; antiparasitic; arteriosclerotic; vulnery; cyostatic;
KW haemopoietic; haematologic; anaemia; autoimmune disorder;
KW rheumatoid arthritis; inflammation; Grave's disease; diabetes;
KW systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
KW Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
KW cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
KW human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003038063-A2.
XX
PD 08-MAY-2003.
XX
PF 19-MAR-2002; 2002WO-US008277.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM,
XX
WP1, 2003-430516/40.
XX
DR P-PSDB; ADC74031.
XX
PT New human secreted polypeptide for diagnosing, preventing or treating
PT hematopoietic or hematologic disorders (e.g. anemia), autoimmune
PT disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
PT atherosclerosis).
XX
PS Claim 27; SEQ ID NO 49; 2272pp; English.
XX
CC The invention relates to a novel human secreted polypeptide comprising a
CC defined sequence given in the specification. The polypeptide, nucleic
CC acid molecule, antibody, agonist or antagonist of the invention may be
CC useful for preparing a composition for diagnosing or treating a
CC haemopoietic or haematologic disorder such as anaemia, autoimmune
CC disorders such as rheumatoid arthritis, inflammation, Grave's disease,
CC diabetes, systemic lupus erythematosus or glomerulonephritis,
CC neurodegenerative disorders including Parkinson's disease and Alzheimer's
CC disease, wounds and hyperproliferative disorders including
CC atherosclerosis or cancer, as well as bacterial, viral, fungal or
CC parasitic infections. The polypeptide may also be used, during gene
CC therapy procedures and for identifying a binding partner by contacting
CC the polypeptide with a binding partner and determining whether the
CC binding partner increases or decreases the activity of the polypeptide.
CC The current sequence is that of the human secreted protein-related DNA of
CC the invention.
XX
SQ Sequence 1038 BP; 328 A; 167 C; 198 G; 339 T; 0 U; 6 Other;
XX
Query Match 99.7%; Score 1035; DB 10; Length 1038;
Best Local Similarity 100.0%; Pred. No. 4,2e-213;
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GNAATTCGACAGAGTAAATGTAATAAATAATTTCTATTAATGAATTTTAATGGAATTAGA 60
DB 1 GNAATTCGACAGAGTAAATGTAATAAATAATTTCTATTAATGAATTTTAATGGAATTAGA 60
QY 61 GCATCATGAGAAAAATGCTCTTACTGTGAAAAACATATTTGTACATTTTGTCAACTA 120
DB 61 GCATCATGAGAAAAATGCTCTTACTGTGAAAAACATATTTGTACATTTTGTCAACTA 120
```



```
QY 121 ATCTTCATTAATTCTTTAGTAACTATAATGTTAAAGTTGACAGGCGAGCTTATATAG 180
D 121 ATCTTCATTAATTCTTTAGTAACTATAATGTTAAAGTTGACAGGCGAGCTTATATAG 180
QY 181 TAAATGCGAGCTGACAGCATGAAATTAATATATATTTTGTGACTATCTTATAGG 240
D 181 TAAATGCGAGCTGACAGCATGAAATTAATATATATTTTGTGACTATCTTATAGG 240
QY 241 AAAATCGAGAAATTCAAAACCTTGTATGTTTAAAGGATATGACATTTTATTAATGT 300
D 241 AAAATCGAGAAATTCAAAACCTTGTATGTTTAAAGGATATGACATTTTATTAATGT 300
QY 301 GCGGATATTTATATACATGATTTGACGTTTGGAATAATTTTCCCTGGACCTTATTTTA 360
D 301 GCGGATATTTATATACATGATTTGACGTTTGGAATAATTTTCCCTGGACCTTATTTTA 360
QY 361 GATGAGATCTACAGTGTAGGCAAACTTATATATCTGTCACTCCATTAGTGTATAGTC 420
D 361 GATGAGATCTACAGTGTAGGCAAACTTATATATCTGTCACTCCATTAGTGTATAGTC 420
QY 421 AGACTCATCCCATGCTTAAATTTATAGTTGKAAAATACGCTTTGTAAATAGTTGTT 480
D 421 AGACTCATCCCATGCTTAAATTTATAGTTGKAAAATACGCTTTGTAAATAGTTGTT 480
QY 481 AGGTCATTATCACCAAGCTTCAAGKATTAATTATTAATAAAGCTTGKTTTATCTTG 540
D 481 AGGTCATTATCACCAAGCTTCAAGKATTAATTATTAATAAAGCTTGKTTTATCTTG 540
QY 541 TGAATACCGTTTTTTCATGCAAAAGTTAAATCTTCAGCCTTTAAATTTTATTAAT 600
D 541 TGAATACCGTTTTTTCATGCAAAAGTTAAATCTTCAGCCTTTAAATTTTATTAAT 600
QY 601 ATATAGAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
D 601 ATATAGAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 GCTTTGCTTAAATGTTTACCGTCAAAATCTTTAGCCAGTTCTTATTTTGTGAT 720
D 661 GCTTTGCTTAAATGTTTACCGTCAAAATCTTTAGCCAGTTCTTATTTTGTGAT 720
QY 721 GAAGATATCTAGTTTATGTCAGAGGCGTGGCGGATAGTGTGATGATGATGATG 780
D 721 GAAGATATCTAGTTTATGTCAGAGGCGTGGCGGATAGTGTGATGATGATGATG 780
QY 781 CTTTGGCGGCGCGAGGAGTGTGATCTTAAGTCAAGAGTTTGAAGACCAAGCTGCCA 840
D 781 CTTTGGCGGCGCGAGGAGTGTGATCTTAAGTCAAGAGTTTGAAGACCAAGCTGCCA 840
QY 841 ACATGCGAAGCGTTGCTCTACTAATAAATTAATAAATTAATAAATTAATAAATTA 900
D 841 ACATGCGAAGCGTTGCTCTACTAATAAATTAATAAATTAATAAATTAATAAATTA 900
QY 901 TCTGTATTTCAAGTACTACAGAGGCTAACAAGAAATTTCTTGAACCTGGAGGAG 960
D 901 TCTGTATTTCAAGTACTACAGAGGCTAACAAGAAATTTCTTGAACCTGGAGGAG 960
QY 961 AGGTGAGTGAAGGCAATTCGATTCAGCCTGGGCAACAAGTGAAGTCTTGTCTAAA 1020
D 961 AGGTGAGTGAAGGCAATTCGATTCAGCCTGGGCAACAAGTGAAGTCTTGTCTAAA 1020
QY 1021 AAAAAAAAAAACTCGA 1038
D 1021 AAAAAAAAAAACTCGA 1038
```

RESULT 7

ABD32800 standard, DNA, 68196 BP.

XX ABD32800;

DT 18-NOV-2004 (first entry)

```
XX Human cancer-associated genomic DNA HD16-049.
XX Human; de; cancer-associated protein; gene; cytostratic; cancer;
XX leukaemia; lymphoma; CAP.
XX Homo sapiens.
XX WO2004074320-A2.
XX 02-SEP-2004.
XX 17-FEB-2004; 2004WO-US004730.
XX 14-FEB-2003; 2003US-00367094.
XX 14-MAR-2003; 2003US-0038838.
XX 15-APR-2003; 2003US-00417375.
XX 13-JUN-2003; 2003US-00461862.
XX 15-SEP-2003; 2003US-00663431.
XX 15-DEC-2003; 2003US-00737318.
XX (SAGR-) SAGRES DISCOVERY INC.
XX Morris DW, Morris DW, Malandro MS;
XX WPI; 2004-652914/63.
XX New isolated cancer-associated polynucleotides and polypeptides useful
XX for diagnosing, preventing or treating cancer, especially lymphoma and
XX leukemia, or in screening for agents that modulate cancer.
XX claim 16; seqid 434; 310pp; English.
XX The invention relates to an isolated nucleic acid comprising at least 10
XX contiguous nucleotides of any of the 233 polynucleotide sequences given
XX in the specification, or its complement. The nucleic acids encode cancer-
XX associated proteins. Also included are an expression vector comprising
XX the isolated nucleic acid cited above, a host cell comprising the above
XX recombinant nucleic acid or expression vector, a microarray for detecting
XX a cancer-associated (CA) nucleic acid comprising at least one probe
XX comprising at least 10 contiguous nucleotides of any of the above-
XX mentioned nucleotide sequences, an isolated polypeptide (encoded within
XX an open reading frame of a CA sequence selected from any of the 95
XX polynucleotide sequences as mentioned in the specification, or its
XX complement), an isolated antibody, (or its antigen binding fragment) that
XX binds to the above polypeptide, a hybridoma that produces the above
XX monoclonal antibody, a pharmaceutical composition comprising the above
XX antibody and a pharmaceutical excipient, a kit for detecting cancer
XX cells/comprising the antibody cited above, methods for diagnosing cancer
XX or for detecting the presence or absence of cancer cells in an
XX individual, a method for inhibiting growth of cancer cells in an
XX individual, a method for delivering a therapeutic agent to cancer cells
XX in an individual, an electronic library comprising the above
XX polynucleotide or polypeptide (or their fragments), methods of screening
XX for anticancer activity or for a bioactive agent capable of modulating
XX the activity of a CA protein (CAP), methods for detecting cancer
XX associated with expression of a polypeptide in a test cell sample, a
XX method for treating cancers and a method for inhibiting the expression of
XX CA gene in a cell. The composition and methods are useful for detecting,
XX diagnosing, preventing and treating cancers, especially lymphoma and
XX leukemia. These may also be used in screening for agents that modulate
XX cancer. The present sequence is a human CAP genomic sequence. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
```

SQ Sequence 68196 BP; 20295 A; 14326 C; 13150 G; 20385 T; 0 U; 40 Other;

Query Match 19.3%; Score 200; DB 13; Length 68196;

Best Local Similarity 81.7%; Pred. No. 4.9e-33;

Matches 245; Conservative 0; Mismatches 50; Indels 5; Gaps 1;

QY 732 GTTTTATGTCAGGGGCTGGCGGATAGTGAAGCTGTGTGCTCCAGTGTCTTGGCGGC 791

PT	variation	32920	/*tag= j	/*tag= ah
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	33929	/*tag= k	/*tag= ai
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	35599	/*tag= l	/*tag= aj
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	36101	/*tag= m	/*tag= ak
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	36340	/*tag= n	/*tag= al
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	36405	/*tag= o	/*tag= am
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	36517	/*tag= p	/*tag= an
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	36777	/*tag= q	/*tag= ao
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	36992	/*tag= r	/*tag= ap
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	37645	/*tag= s	/*tag= aq
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	37868	/*tag= t	/*tag= ar
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	38440	/*tag= u	/*tag= as
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	38532	/*tag= v	/*tag= at
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	38547	/*tag= w	/*tag= au
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	38712	/*tag= x	/*tag= av
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	40684	/*tag= y	/*tag= aw
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	40860	/*tag= z	/*tag= ax
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	41213	/*tag= aa	/*tag= ay
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	41419	/*tag= ab	/*tag= az
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	41613	/*tag= ac	/*tag= ba
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	42407	/*tag= ad	/*tag= bb
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	43440	/*tag= ae	/*tag= bc
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	44247	/*tag= af	/*tag= bd
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	44677	/*tag= ag	/*tag= be
PT	/standard_name=	"Single nucleotide polymorphism (SNP)"	/standard_name=	"Single nucleotide polymorphism (SNP)"
PT	variation	45256	/*tag= bf	

PD 02-SEP-2004.
XX
PR 17-FEB-2004; 2004MO-US004730.
XX
PR 14-FEB-2003; 2003US-00367094.
PR 14-MAR-2003; 2003US-00388838.
PR 15-APR-2003; 2003US-00417375.
PR 13-JUN-2003; 2003US-00461862.
PR 15-SEP-2003; 2003US-00663431.
PR 15-DEC-2003; 2003US-00737318.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morrie DW, Morris DW, Malandro MS;
XX WPI; 2004-652914/63.
XX
PT Now isolated cancer-associated polynucleotides and polypeptides useful
PT for diagnosing, preventing or treating cancers, especially lymphoma and
PT leukemia, or in screening for agents that modulate cancer.
XX
PS claim 16, seqid 24; 310pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising at least 10
CC contiguous nucleotides of any of the 233 polynucleotide sequences given
CC in the specification, or its complement. The nucleic acids encode cancer-
CC associated proteins. Also included are an expression vector comprising
CC the isolated nucleic acid cited above, a host cell comprising the above
CC recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells (comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a human CAP genomic sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 310268 BP; 87522 A; 60932 C; 62901 G; 98913 T; 0 U; 0 Other;

QY 925 GCTAACACAGAAAAATTCTTGAACCTGGAGGACAGAGTTGCAGTAG----- 973
DB 140297 GCTAAGGACAGAAATGCTTGAACCTGGAGGACAGAGTTGCAGTAGCCCAATCACA 140238
QY 974 CCAATTGACCTCCAGCCTGGGCAACACAGTAGACTCTTGTCTCAAAAAAAAAAAAAA 1033
DB 140237 CCACTGACCTCCAGCCTGGGCAACAGAGTAGACTCTGCTCACTTAATAAAAAAAAAA 140178
RESULT 12
AAL04931
ID AAL04931 standard; DNA; 32249 BP.
XX
AC AAL04931;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 7619.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; db.
XX
XX Homo sapiens.
OS
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218250P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225265P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225475P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226888P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227709P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.

[illegible]

PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244E.
PR	17-NOV-2000;	2000US-0249245F.
PR	17-NOV-2000;	2000US-0249264D.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391E.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	06-DEC-2000;	2000US-0256719P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
DR	WPI; 2001-465570/50.	
XX		
PT	Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition.	
XX		
PS	Disclosure; SEQ ID NO 7619; 1297bp + Sequence Listing; English.	
XX		
CC	The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a genomic sequence encoding a protein of the invention	
CC		
CC		
SQ	Sequence 32249 BP; 9657 A; 6826 C; 7366 G; 8400 T; 0 U; 0 Other;	
Query Match	19.1%; Score 198.2; DB 4; Length 32249;	
Best Local Similarity	81.9%; Pred. NO. 1e-32; Indels 11; Gaps 1;	
Matches 245; Conservative 0; Mismatches 43;		
DB	745 GGCTGGGCGCATAGTCATGCCCTGTGATCCCAAGTCTTTCCGGGCGCAGGACGTGA 804 11877 GGCTGGGCGCGTGTCATGCCTGTATCCCAACTTTGGGAGGCCGAGTGGGACAGA 11933	
OY	805 TCACCTTAAGTCAGAGATTGAGACCCAGCCTGCCCAACATGTGGAACGTTGCTCTACT 864 11937 TCACCTTAGGTCAGAGATTGAGACCCAGCCTGGCCAACATGAGTGAACCCGCTTCTACT 11999	
DB	865 AAAAATACAAAATAATTAGACAGAGCGTGTGGCACACATCTGTAATTCCAGCTACTCAGAG 924 11997 AAAAATACAAAATAATTAGACAGCGCTGTGTGTCATTAATCCAGCTACTCAGAG 12055	
OY	925 GCTAACACAGAAAAATTCCTTGAACCTGGAGGACAGAGTTGCAGTGAGCCAT----- 977 12057 GCTAGACAGAGAGATGCTTTGAACCTGGAGGACAGAGTGTCTGTGAGCCATGACTGCA 12116	
OY	978 -----TGACATCCAGCCCTGGGCAACACAGTGAAGCTGTGTCTCAAAAAAAAAAAAAAA 1032 12117 CCACCTGACATCCAGCCCTGGGCAACAGAGTGAAGCTGTCTCAAAAAAAAAAAAAAGAA 12175	
RESULT 13		
ABL97825		
ID	ABL97825 standard; DNA; 32249 BP.	
AC	ABL97825;	
DT	21-JUN-2002 (first entry)	

XX Human testicular antigen encoding DNA fragment SEQ ID NO: 2477.
DE
XX
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS Homo sapiens.
XX
PN W0200155317-A2.
XX
PD 02-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US001329.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 26-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232387P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 25-SEP-2000; 2000US-0234948P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0249207P.
PR 08-NOV-2000; 2000US-0249208P.
PR 08-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249259P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2005, 04:14:19 ; Search time 220 Seconds
(without alignments)
7720.251 Million cell updates/sec

Title: US-10-664-358-36
Perfect score: 1038
Sequence: 1 gnaatcgcgcagcagctaac.....aaaaaaaaaaaaaacctcga 1038

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUS.COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1011.6	97.5	168104	US-09-949-016-12026	Sequence 15026, A
2	1011.6	97.5	168105	US-09-949-016-16554	Sequence 16554, A
3	598.6	57.7	601	US-09-949-016-170431	Sequence 170431, A
4	200	19.3	51620	US-09-949-016-12848	Sequence 12848, A
5	200	19.3	51621	US-09-949-016-15503	Sequence 15503, A
6	200	19.3	75799	US-09-949-016-15231	Sequence 15231, A
7	198.8	19.2	14345	US-09-949-016-15449	Sequence 15449, A
8	198.8	19.2	39243	US-09-949-016-12316	Sequence 12316, A
9	198.8	19.2	39243	US-09-949-016-15443	Sequence 15443, A
10	198.4	19.1	6201	US-09-949-016-131748	Sequence 131748, A
11	198.2	19.1	88490	US-09-949-016-12758	Sequence 12758, A
12	198.2	19.1	88736	US-09-949-016-14222	Sequence 14222, A
13	196.8	19.0	601	US-09-949-016-15966	Sequence 15966, A
14	196.8	19.0	20441	US-09-949-016-13052	Sequence 13052, A
15	196.4	18.9	601	US-09-949-016-15905	Sequence 15905, A
16	196.4	18.9	89689	US-09-949-016-13089	Sequence 13089, A
17	196.2	18.9	601	US-09-949-016-121170	Sequence 121170, A
18	196.2	18.9	601	US-09-949-016-127095	Sequence 127095, A
19	196.2	18.9	601	US-09-949-016-140199	Sequence 140199, A
20	196.2	18.9	71119	US-09-949-016-15358	Sequence 15358, A
21	196	18.9	73853	US-09-949-016-12029	Sequence 12029, A
22	195.8	18.9	601	US-09-949-016-127094	Sequence 127094, A
23	195.6	18.8	90618	US-09-949-016-15964	Sequence 15964, A
24	195.4	18.8	17132	US-09-949-016-15361	Sequence 15361, A
25	195.4	18.8	24150	US-09-949-016-12438	Sequence 12438, A
26	195.2	18.8	45427	US-09-949-016-16243	Sequence 16243, A
27	195	18.8	601	US-09-949-016-15626	Sequence 15626, A

C	28	195	18.8	18200	4	US-09-949-016-15660	Sequence 15660, A
C	29	195	18.8	18200	4	US-09-949-016-15661	Sequence 15661, A
C	30	195	18.8	19719	4	US-09-949-016-15662	Sequence 15662, A
C	31	195	18.8	19719	4	US-09-949-016-15663	Sequence 15663, A
C	32	195	18.8	55674	4	US-09-949-016-12563	Sequence 12563, A
C	33	195	18.8	55675	4	US-09-949-016-15706	Sequence 15706, A
C	34	195	18.8	74804	4	US-09-949-016-15118	Sequence 15118, A
C	35	194.6	18.7	601	4	US-09-949-016-140198	Sequence 140198, A
C	36	194.6	18.7	601	4	US-09-949-016-140200	Sequence 140200, A
C	37	194.6	18.7	219964	4	US-09-949-016-15086	Sequence 15086, A
C	38	194.4	18.7	601	4	US-09-949-016-63026	Sequence 63026, A
C	39	194.4	18.7	601	4	US-09-949-016-63027	Sequence 63027, A
C	40	194.4	18.7	58543	4	US-09-949-016-13565	Sequence 13565, A
C	41	194.4	18.7	373182	4	US-09-949-016-17371	Sequence 17371, A
C	42	194.4	18.7	373694	4	US-09-949-016-12062	Sequence 12062, A
C	43	194.4	18.7	678533	4	US-09-949-016-14577	Sequence 14577, A
C	44	194.4	18.7	678533	4	US-09-949-016-14578	Sequence 14578, A
C	45	194	18.7	25401	4	US-09-949-016-13345	Sequence 13345, A

ALIGNMENTS

```
RESULT 1
US-09-949-016-12026
Sequence 12026, Application US/09949016
Patent NO. 681239
GENERAL INFORMATION:
APPLICANT: VENTUR, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12026
LENGTH: 168104
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1) .. (168104)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12026
Query Match 97.5%; Score 1011.6; DB 4; Length 168104;
Best Local Similarity 99.1%; Pred. No. 8.5e-237;
Matches 1011; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
16 TTAATGATTAATTAATTTCTTAATGAAATTTTAATGGAATTAAGCATCATGAAAAA 75
82814 TTAATGATTAATTAATTTCTTAATGAAATTTTAATGGAATTAAGCATCATGAAAAA 82873
76 TGCCTTACGTGGAACATTTATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTT 82933
82874 TGCCTTACGTGGAACATTTATTTGTCATTTGTCATTTGTCATTTGTCATTTGTCATTT 82933
136 TTAATGATTAATTAATTTCTTAATGAAATTTTAATGGAATTAAGCATCATGAAAAA 195
82934 TTAATGATTAATTAATTTCTTAATGAAATTTTAATGGAATTAAGCATCATGAAAAA 82993
196 AGCATGAAATTAATTTCTTAATGAAATTTTAATGGAATTAAGCATCATGAAAAA 255
82994 AGCATGAAATTAATTTCTTAATGAAATTTTAATGGAATTAAGCATCATGAAAAA 83053
256 CAAACCTGTTAGTTTATGAGGATAGTCACATTTATTAATGCGGTATATTATAC 315
```

Db	83054	CAAAACCTGTAGTCTTTAGGGTATAGTCACTTTATPAATGTCCGTAATTTATAC	83113
Oy	316	ATGATTTGACGTTTGTGMAAATAATTTCCCTGGACCTTTATTTTAAAGTACTTACAGT	375
Db	83114	ATGATTTGACGTTTGTGMAAATAATTTCCCTGGACCTTTATTTTAAAGTACTTACAGT	83173
Oy	376	GTAGGCAACTTATATATACTGTGCACCTCCATTAGTGCATAGTCAGACTCATCCCATG	435
Db	83174	GTAGGCAACTTATATATATCTGCACCTCCATTAGTGCATAGTCAGACTCATCCCATG	83233
Oy	436	CTAAATTAATAGTTGTGAATAATACGCTTTGTGAATAAGTGTGTAGTCATTATCAACA	495
Db	83234	CTAAATTAATAGTTGTGAATAATACGCTTTGTGAATAAGTGTGTAGTCATTATCAACA	83293
Oy	496	AGTCTTCAGGKATTCATTATATATAAACCCTTGGKTTTATTCCTGTGAATMCCGTTTT	555
Db	83294	AGTCTTCAGGATTAATCATATATAAACCCTTGGTTTTTATTCCTGTGAATAACCGTTTT	83353
Oy	556	TCCATGCAAGTTTAAATTCCTCAGCCCTTAATTTTTTTATTAATATATPAAGATGTAT	615
Db	83354	TCCATGCAAGTTTAAATTCCTCAGCCCTTAATTTTTTTATTAATATATPAAGATGTAT	83413
Oy	616	GAGTATGACTACAAAACAGAAAAAATPAACAGATTCGTTGTGGCTTTTGCTAAATTG	675
Db	83414	GAGTATGACTACAAAACAGAAAAAATPAACAGATTCGTTGTGGCTTTTGCTAAATTG	83473
Oy	676	TTACTTGCMAAATCTTAGCCAGTCTTCAATTTTCGTTTGAATGTAACATCTTAAGTT	735
Db	83474	TTACTTGCMAAATCTTAGCCAGTCTTCAATTTTCGTTTGAATGTAACATCTTAAGTT	83533
Oy	736	TAGTCCAGGGGCTGGGCGCATAGCTGATGCTGTGGTCCAGTGCTTTGCGGGCCGAG	795
Db	83534	TAGTCCAGGGGCTGGGCGCATAGCTGATGCTGTGGTCCAGTGCTTTGCGGGGCCAG	83593
Oy	796	GCAGGTGATCATTTAAGTTCAGGAGTTTGAACAACAGCTGCCCAACATGTGAAAAGTT	855
Db	83594	GCAGGTGATCATTTAAGTTCAGGAGTTTGAACAACAGCTGCCCAACATGTGAAAAGTT	83653
Oy	856	GTCCTACTAAAAATACAAAAATTAGA CAGGCGGTGGGCAACATCTTAATTCAGCT	915
Db	83654	GTCCTACTAAAAATACAAAAATTAGA CAGGCGGTGGGCAACATCTTAATTCAGCT	83713
Oy	916	ACTCAGGAGGCTAACACAGGAAAAATTCCTTGAACCTGGGAGGCGAGGTTGCAGTGAGCC	975
Db	83714	ACTCAGGAGGCTAACACAGGAAAAATTCCTTGAACCTGGGAGGCGAGGTTGCAGTGAGCC	83773
Oy	976	ATTGCATTCAGGCTGGGCAACAACAGTGAACCTCTTGTCCTCAAAAAAAAAAAAAAACT	1035
Db	83774	ATTGCATTCAGGCTGGGCAACAACAGTGAACCTCTTGTCCTCAAAAAAAAAAAAAAACT	83833

RESULT 2
 US-09-949-016-16554
 : Sequence 16554, Application US/09949016
 : Patent No. 6812339
 : GENERAL INFORMATION:
 : APPLICANT: VENTER, J. Craig et al.
 : TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 : TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 : FILE REFERENCE: CL001307
 : CURRENT APPLICATION NUMBER: US/09/949,016
 : CURRENT FILING DATE: 2000-04-14
 : PRIOR APPLICATION NUMBER: 60/241,755
 : PRIOR FILING DATE: 2000-10-20
 : PRIOR APPLICATION NUMBER: 60/237,768
 : PRIOR FILING DATE: 2000-10-03
 : PRIOR APPLICATION NUMBER: 60/231,498
 : PRIOR FILING DATE: 2000-09-08
 : NUMBER OF SEQ ID NOS: 207012
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 16554
 : LENGTH: 168105

```

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(168105)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16554

```

Query Match	97.5%	Score 1011.6	DB 4	Length 168105
Best Local Similarity	99.1%	Pred. No. 8.5e-237		
Matches 1011	Conservative 5	Mismatches 4	Indels 0	Gaps 0
QY	16	TTAATGATAAAATATTTCTATATGAAATTTTATATGGGAATTAGACATCATAGAAAAA	75	
Db	82814	TTAATGATAAAATATTTCTATATGAAATTTTATATGGGAATTAGACATCATAGAAAAA	82872	
QY	76	TGCTCTTACGTGTGAAAACATTAATTTGTTACATTTTGGTCACTAATCTTTCAATACTT	135	
Db	82874	TGCTCTTACGTGTGAAAACATTAATTTGTTACATTTTGGTCACTAATCTTTCAATACTT	82933	
QY	136	TTTGTATCTAATATGTTAAAGTTGTACACAGTGGAGCTTAATATAGAAATAGCAGCTGAC	195	
Db	82934	TTTGTATCTAATATGTTAAAGTTGTACACAGTGGAGCTTAATATAGAAATAGCAGCTGAC	82992	
QY	196	AGCATGAAATTAACATATCTAATATTTTGTGACATCTATTATGAAAAACAGAGATT	255	
Db	82994	AGCATGAAATTAACATATCTAATATTTTGTGACATCTATTATGAAAAACAGAGATT	83053	
QY	256	CAAAACCTGTATGTTTATAGGATATAGTCACTTTTATTAATATGCGGTATTTATAC	315	
Db	83054	CAAAACCTGTATGTTTATAGGATATAGTCACTTTTATTAATATGCGGTATTTATAC	83113	
QY	316	ATATTTGAGTGTGTGMAAATATTTTCCCTGACCTTTTATTTAGATGAGATCTACAGT	375	
Db	83114	ATATTTGAGTGTGTGMAAATATTTTCCCTGACCTTTTATTTAGATGAGATCTACAGT	83172	
QY	376	GTAAGCAAACTTATATATATCTGTCAACTCCATTAGTGTGATAGTCAAGTCAATCCCATG	435	
Db	83174	GTAAGCAAACTTATATATATCTGTCAACTCCATTAGTGTGATAGTCAAGTCAATCCCATG	83233	
QY	436	CTAAATTTATAGTTGTAAATATCGCTTTTGTAAATAGTTGTATAGTCAATATACCA	495	
Db	83234	CTAAATTTATAGTTGTAAATATCGCTTTTGTAAATAGTTGTATAGTCAATATACCA	83292	
QY	496	AGTCTTCAAGGKATTAATATATTAATAAACCCTGGKTTTATCTTGTGAATAMCGTTTT	555	
Db	83294	AGTCTTCAAGGKATTAATATATTAATAAACCCTGGKTTTATCTTGTGAATAMCGTTTT	83353	
QY	556	TCCATGCAAAATTAATATCTTCAGCCTTTAATTTTATTAATATATAGATGTGAT	615	
Db	83354	TCCATGCAAAATTAATATCTTCAGCCTTTAATTTTATTAATATATAGATGTGAT	83411	
QY	616	GAGTATGATCTACAAAACAGAAAAATTAACAGATTTTCGTTGTGCTTTGTAAATG	675	
Db	83414	GAGTATGATCTACAAAACAGAAAAATTAACAGATTTTCGTTGTGCTTTGTAAATG	83473	
QY	676	TTTACCTGACAAATCTTATGACAGTCTCTCAATTTTCGTTTGTAGATGAATACTTATGTT	735	
Db	83474	TTTACCTGACAAATCTTATGACAGTCTCTCAATTTTCGTTTGTAGATGAATACTTATGTT	83533	
QY	736	TAGTCCAGGGGCTGGGCGCGATAGCTGATGCTGTGTCCAGTGTCTTTCGGGGCGAG	795	
Db	83534	TAGTCCAGGGGCTGGGCGCGATAGCTGATGCTGTGTCCAGTGTCTTTCGGGGCGAG	83592	
QY	796	GCAAGTGTGATCTTAAAGTCAAGGAGTTTGAAGACAGCTGCGCCCAATATGTGAACGTT	855	
Db	83594	GCAAGTGTGATCTTAAAGTCAAGGAGTTTGAAGACAGCTGCGCCCAATATGTGAACGTT	83653	
QY	856	GTCCTACTAATAAAATCAAAAAATTTACACAGGCGTGTGACACACATCTGTAATTCAGCT	915	
Db	83654	GTCCTACTAATAAAATCAAAAAATTTACACAGGCGTGTGACACACATCTGTAATTCAGCT	83713	
QY	916	ACTCAGAGGCTTAACAAGAAAAATTCCTTGAACCTGGAGGCGAGAGTTTGCAGTGAGCC	975	

Db	Qy	Db
83714	ATTGCACTCCAGCCTGGGCAACAGTGAAGCTCTTGTCACAAAAAAAAAAAAAACT	83773
ACTCAGAGGCTAACACAGGAAAAATTCCTTGAACCTGGAGGCGAGAGTTCCAGTGAAGCC	976	ATTGCACTCCAGCCTGGGCAACAGTGAAGCTCTTGTCACAAAAAAAAAAAAAACT
83774	ATTGCACTCCAGCCTGGGCAACAGTGAAGCTCTTGTCACAAAAAAAAAAAAAAAT	83833

RESULT 3

```

US-09-949-016-170431/c
/ Sequence 170431, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH INFECTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 170431
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-170431

```

Query March	57.7%	Score 598.6;	DB 4	Length 601;
Best Local Similarity	99.0%	Pred. No. 7.8e-137		
Matches 595;	Conservative	6;	Mismatches 0;	Indels 0;
				Gaps 0;

Qy	62	CATCATGAAAAAAATGCTCTTACCTGTGAAAAACATTAATTGTGACATTTGGTCAACCTAA	121
Db	601	CATCATGAAAAAAATGCTCTTACCTGTGAAAAACATTAATTGTGACATTTGGTCAACCTAA	543
Qy	122	TCCTTCATPACTTTTAGTACCTATPANGTAAAGTTGACAGGTGGCAGCTCTTATPAGT	181
Db	541	TCCTTCATPACTTTTAGTACCTATPANGTAAAGTTGACAGGTGGCAGCTCTTATPAGT	482
Qy	182	AAATGGCAGGTGACAGCATGAAAAATPACATCTPAAATTTTGGACATCTTATPAGGA	241
Db	481	AAATGGCAGGTGACAGCATGAAAAATPACATCTPAAATTTTGGACATCTTATPAGGA	422
Qy	242	AAATCAGAGATTTTCAAAAACCTTGTAGTTTTAGGGTATAGTCACATTTATPAAATGTG	301
Db	421	AAATCAGAGATTTTCAAAAACCTTGTAGTTTTAGGGTATAGTCACATTTATPAAATGTG	362
Qy	302	CGGTATATTTATPACATGATTTGACGCTGTGTGAAATATTTTCCCTGGACCTTTATPTTAG	361
Db	361	CGGTATATTTATPACATGATTTGACGCTGTGTGAAATATTTTCCCTGGACCTTTATPTTAG	302
Qy	362	ATGAGATCTACAGGTGAGCAAACTTATATPACTGTCAACTCATTAAGTGTCAATGTCA	421
Db	301	ATGAGATCTACAGGTGAGCAAACTTATATPACTGTCAACTCATTAAGTGTCAATGTCA	242
Qy	422	GACTCATCCCGCAGTCPAAAAATPANGTGTGKAAAATAGCCTTTGTGAAATAGTGTGTGA	481
Db	241	GACTCATCCCGCAGTCPAAAAATPANGTGTGKAAAATAGCCTTTGTGAAATAGTGTGTGA	182
Qy	482	GGTCATTAATCCCAAGCTCTTCAAGGATATACATTAATAAAAACCTTGGKTTTATTTCTGT	541
Db	181	GGTCATTAATCCCAAGCTCTTCAAGGATATACATTAATAAAAACCTTGGTTTATTTCTGT	122
Qy	542	GAATPACCGTTTTTTCATATGCAAAAGTTAAATTTCTTACGCTTTAATTTTATPATA	601
Db	121	GAATPACCGTTTTTTCATATGCAAAAGTTAAATTTCTTACGCTTTAATTTTATPATA	62

Oy 602 TATAAGATGTGATGATGACTACAAACAGGAAAAAATAACAGATTTCGTTGTGG 661
 Db 61 TATAAGATGTGATGATGACTACAAACAGGAAAAAATAACAGATTTCGTTGTGG 2
 Oy 662 C 662
 Db 1 C 1

RESULT 4

```

US-09-949-016-12848
; Sequence 12848, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01107
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12848
; LENGTH: 51620
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(51620)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12848

```

Query Match	19.3%	Score 200;	DB 4;	Length 51620;
Best Local Similarity	81.7%	Pred. No. 7.4e-39;		
Matches 245; Conservative	0;	Mismatches 50;	Indels 5;	Gaps 1;

Qy	732	GTTTTAGTCACAGGGCTGGGCGCATACTGATGCTGGAGTCCACAGTCTTTGCGGGC	791
Db	8293	GTATTACCTTTAAGCTGGGCAATGCTGATGCTGTAACTCCCTGCACTTTGGAGGC	898
Qy	792	CGAGGACAGTGGATCTCTTAAAGTCAGGAGTTTGAGCCAGCCTTGCCCAACATGTGAAA	851
Db	8989	CGAGGACAGTGGATCTCTTAAAGTCAGGAGTTCAAGACACAGCCTGGCCAAACATGTGAAA	904
Qy	852	CGTTGTCTCTACTTAAATAACAAAAATTGACAGGCGTGGTGGCACACATCTGTATTC	911
Db	9049	CTGGTCTCTACTTAAATAACAAAAATTGACAGGACATGTTGTGTGACCTGTAAATCC	910
Qy	912	AGTACTCAGAGGCTTAACACAGSAAAAATTCCTTGAACCTGGGAGGACAGAGTTCGACGTG	971
Db	9109	AGTACTCAGAGGCTGAGGACAGGAGAAATTTGCTTGAACCTGGGAGGCGGAG-----ATTG	916
Qy	972	AGCCATTGCAATCCAGCCTGGGCAACACAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAA	103
Db	9164	TGCCAAAGCACTCCAGCCTGGGCAACAAATATGACTTTGTGTCCAAAAAAAAAAAAAAAAA	922

RESULT 5

US-09-949-016-16503
Sequence 16503, Application US/05949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01107
CURRENT APPLICATION NUMBER: US/09/949-016

```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 16503
; LENGTH: 51621
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(51621)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16503
```

```

Query Match          19.3%; Score 200; DB 4; Length 51621;
Best Local Similarity 81.7%; Pred. No. 7,4e-39;
Matches 245; Conservative 0; Mismatches 50; Indels 5; Gaps 1;
```

```

QY 732 GTTTAGTTCAGGGGCTGGGCGGATGCTGATGCTGTGCTCCAGTCTTTGCGGGC 791
    |||
DB 8929 GTATTATCCCTTAGGCTGGGCGACATGCTGATGCTGTAACTCTGACATTTGGAGGC 8988
    |||
QY 792 CGAGCAGGTGGATCACTTAAGTCAGAGGATTGAGCAGCCTGCCCAACATGTGAAA 851
    |||
DB 8989 CGAGCAGGTGGATCACTTAAGTCAGAGGATTGAGCAGCCTGCCCAACATGTGAAA 9048
    |||
QY 852 CGTGTCTCTACTAATAAAATACAAAATTAGACAGGCGTGTGGCACAATCTGTAATTC 911
    |||
DB 9049 CTTGTCTCTACTAATAAAATACAAAATTAGACAGGCGTGTGGCACAATCTGTAATTC 9108
    |||
QY 912 AGCTACTAGAGAGCTTAACACAGAAAATTCCTTGAACCTTGGAGGCGAGGTTGCAGTG 971
    |||
DB 9109 AGCTACTAGAGAGCTTAACACAGAAAATTCCTTGAACCTTGGAGGCGAGGTTGCAGTG 9163
    |||
QY 972 AGCCATTGCACTCCAGCCTGGGCAACACAGTGAAGCTTTGTCCAAAAAATTTGCAAAA 1031
    |||
DB 9164 TGCCAAAGCACTCCAGCCTGGGCAACAAATGAGACTTTGTGTCCAAAAAATTTGCAAAA 9223
    |||
```

```

RESULT 6
US-09-949-016-15231/c
; Sequence 15231, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 15231
; LENGTH: 75799
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15231
```

```

Query Match          19.3%; Score 200; DB 4; Length 75799;
Best Local Similarity 81.7%; Pred. No. 8,4e-39;
Matches 245; Conservative 0; Mismatches 50; Indels 5; Gaps 1;
```

```

QY 732 GTTTAGTTCAGGGGCTGGGCGGATGCTGATGCTGTGCTCCAGTCTTTGCGGGC 791
    |||
DB 9102 GTATTATCCCTTAGGCTGGGCGACATGCTGATGCTGTAACTCTGACATTTGGAGGC 9043
    |||
QY 792 CGAGCAGGTGGATCACTTAAGTCAGAGGATTGAGCAGCCTGCCCAACATGTGAAA 851
    |||
DB 9042 CGAGCAGGTGGATCACTTAAGTCAGAGGATTGAGCAGCCTGCCCAACATGTGAAA 8983
    |||
QY 852 CGTGTCTCTACTAATAAAATACAAAATTAGACAGGCGTGTGGCACAATCTGTAATTC 911
    |||
DB 8982 CTTGTCTCTACTAATAAAATACAAAATTAGACAGGCGTGTGGCACAATCTGTAATTC 8923
    |||
QY 912 AGCTACTAGAGAGCTTAACACAGAAAATTCCTTGAACCTTGGAGGCGAGGTTGCAGTG 971
    |||
DB 8922 AGCTACTAGAGAGCTTAACACAGAAAATTCCTTGAACCTTGGAGGCGAGGTTGCAGTG 8868
    |||
QY 972 AGCCATTGCACTCCAGCCTGGGCAACACAGTGAAGCTTTGTCCAAAAAATTTGCAAAA 1031
    |||
DB 8867 TGCCAAAGCACTCCAGCCTGGGCAACAAATGAGACTTTGTGTCCAAAAAATTTGCAAAA 8808
    |||
```

```

RESULT 7
US-09-949-016-15449/c
; Sequence 15449, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 15449
; LENGTH: 14345
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15449
```

```

Query Match          19.2%; Score 198.8; DB 4; Length 14345;
Best Local Similarity 76.9%; Pred. No. 9,2e-39;
Matches 260; Conservative 0; Mismatches 67; Indels 11; Gaps 1;
```

```

QY 707 TTTCTTTGAGATGATAATTAAGTTTATGTCAGGGGCTGGGCGGATGATGATGC 766
    |||
DB 9994 TTAAGTTTCTTTGTTAAGCATTTTGAATTTGAGAAATGGCTGGGCGGATGATGC 9935
    |||
QY 767 CTGTGTCCAGTGTCTTTCGCGGCGGAGCAGGTGATCACTTAAGTCAAGATTTGA 826
    |||
DB 9934 CTGTATCCAGCACTTTGGAGGCGGAGGTGGATCACTTAAGTCAAGATTTGA 9875
    |||
QY 827 GACGAGCCTGCCCAACAGTGTGAAAATCTTGTCTCTACTAATAAAATTTGACAGG 886
    |||
DB 9874 GACGAGCCTGCCCAACAGTGTGAAAATCTTGTCTCTACTAATAAAATTTGACAGG 9815
    |||
QY 887 CGTGTGACACATCTGTAATTCAGTCACTGAGAGGCTTAACACAGAAAATTTCCGTG 946
    |||
DB 9814 CATGTGACATGACCTGTATGCTTCCAGTCACTTGGAGGCTGAGGAGAAATTCGCTTG 9755
    |||
QY 947 AACCTGAGAGGAGGTTGAGTAG-----CCATTGCACTCCAGCTGGGCA 995
    |||
DB 9754 AACCCAGAGGCGGAGGTTGCAATGAGCAAGATTGCAACATTCACCTCCAGCTGGGTG 9695
    |||
QY 996 AACAGTGAGACTCTTGTCTCAAAAAAATTTGCAAAA 1033
    |||
DB 9694 ACAAGCGAGACTCATCTCAAAAAAATTTGCAAAA 9657
    |||
```

RESULT 8

US-09-949-016-12316
Sequence 12316, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12316
LENGTH: 39243
TYPE: DNA
ORGANISM: Human
US-09-949-016-12316

Query Match 19.2%; Score 198.8; DB 4; Length 39243;
Best Local Similarity 76.9%; Pred. No. 1.3e-38;

Matches 260; Conservative 0; Mismatches 67; Indels 11; Gaps 1;

Qy 707 TTTCGTTTGAGATGAGATCTAGTTTATGTCAGAGGCGCGAGTAGCTGATGC 766
Db 38368 TTAAGTTTCTGTTAAGCATTTTAGAATTGAGAAATAGGTGGCGGCTGCTACGC 38427
Qy 767 CTGTGTCACAGTCTTTGGCGGCGCGAGCGAGTGATCATCTTAAGTGA 826
Db 38428 CTGTATCCAGACCTTTGGAGGCGGAGGTGGTGTATCATCTGAGTCAAGTTGA 38487
Qy 827 GACCACTGCGCCCAACATGCTGTAACCTTCTCTACTAAATAATTAAGACAG 886
Db 38488 GACCACTGCGCCCAACATGCTGTAACCTTCTCTACTAAATAATTAAGACAG 38547
Qy 887 CGTGTGCGACACATCTGTATTCAGCTACTCAGAGGCTAACAAGAAATTCCTTG 946
Db 38548 CATGTGCGACACCTGTGTATTCAGCTACTCAGAGGCTAACAAGAAATTCCTTG 38607
Qy 947 AACCTGGAGGAGAGGTTGACAGTGA-----CCATTGCACTCCAGCTGGCA 995
Db 38608 AACCTGGAGGAGGTTGACAGTGA-----CCATTGCACTCCAGCTGGCA 995
Qy 996 ACAAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 1033
Db 38608 ACAAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 38705

RESULT 9

US-09-949-016-15443
Sequence 15443, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15443
LENGTH: 39243
TYPE: DNA
ORGANISM: Human
US-09-949-016-15443

Query Match 19.2%; Score 198.8; DB 4; Length 39243;
Best Local Similarity 76.9%; Pred. No. 1.3e-38;

Matches 260; Conservative 0; Mismatches 67; Indels 11; Gaps 1;

Qy 707 TTTCGTTTGAGATGAGATCTAGTTTATGTCAGAGGCGCGAGTAGCTGATGC 766
Db 38368 TTAAGTTTCTGTTAAGCATTTTAGAATTGAGAAATAGGTGGCGGCTGCTACGC 38427
Qy 767 CTGTGTCACAGTCTTTGGCGGCGCGAGCGAGTGATCATCTTAAGTGA 826
Db 38428 CTGTATCCAGACCTTTGGAGGCGGAGGTGGTGTATCATCTGAGTCAAGTTGA 38487
Qy 827 GACCACTGCGCCCAACATGCTGTAACCTTCTCTACTAAATAATTAAGACAG 886
Db 38488 GACCACTGCGCCCAACATGCTGTAACCTTCTCTACTAAATAATTAAGACAG 38547
Qy 887 CGTGTGCGACACATCTGTATTCAGCTACTCAGAGGCTAACAAGAAATTCCTTG 946
Db 38548 CATGTGCGACACCTGTGTATTCAGCTACTCAGAGGCTAACAAGAAATTCCTTG 38607
Qy 947 AACCTGGAGGAGAGGTTGACAGTGA-----CCATTGCACTCCAGCTGGCA 995
Db 38608 AACCTGGAGGAGGTTGACAGTGA-----CCATTGCACTCCAGCTGGCA 995
Qy 996 ACAAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 1033
Db 38608 ACAAGTGAAGCTTGTCTCAAAAAAAAAAAAAAAAAAAAA 38705

RESULT 10

US-09-949-016-131748
Sequence 131748, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 131748
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-131748

Query Match 19.1%; Score 198.4; DB 4; Length 601;
Best Local Similarity 76.6%; Pred. No. 3.8e-39;

Matches 259; Conservative 1; Mismatches 67; Indels 11; Gaps 1;

Qy 707 TTTCGTTTGAGATGAGATCTAGTTTATGTCAGAGGCGCGAGTAGCTGATGC 766
Db 256 TTAAGTTTCTGTTAAGCATTTTAGAATTGAGAAATAGGTGGCGGCTGCTACGC 315
Qy 767 CTGTGTCACAGTCTTTGGCGGCGCGAGCGAGTGATCATCTTAAGTGA 826
Db 316 CTGTATCCAGACCTTTGGAGGCGGAGGTGGTGTATCATCTGAGTCAAGTTGA 375

QY 827 GACAGCCTGCGCAACATGTTGAAACGTTGCTCTACTATAAATAATTAACAGAG 886
| | | | |
Db 376 GACCAAGCTGGCAACATGTTGAAACCTTGTCTCTATAAATAATTAACAGAG 435
| | | | |
QY 887 CGTGTGGCAACATCTGTATTCAGCTACTAGAGAGGCTTAACAGAAATTCCTTG 946
| | | | |
Db 436 CATGTGGCATGACCTGTATGCTCCAGCTACTTGGAGGCTGAGAGAAATGCTTG 495
| | | | |
QY 947 AACCTGGAGAGAGAGTTGCAATGAG-----CCATTGCACTCCAGCTGGAGCA 995
| | | | |
Db 496 AACCCAGAGAGCGAGGTGCAATGAGCCAAAGATTCACCACTGCACTCCAGCTGGAG 555
| | | | |
QY 996 ACACAGTGAAGCTCTGTCTCAAAAAAAAAAAAAAAAAAAAA 1033
| | | | |
Db 556 ACAGAGGAGAGCTCCATCTCAAAAAAAAAAAAAAAAAAAAA 593
| | | | |

RESULT 11
US-09-949-016-12758
; Sequence 12758, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12758
; LENGTH: 88490
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12758

Query Match 19.1%; Score 198.2; DB 4; Length 88490;
Best Local Similarity 81.9%; Pred. No. 2.4e-38;
Matches 245; Conservative 0; Mismatches 43; Indels 11; Gaps 1;
QY 745 GAGTGGGCGCATAGCTGATGCTGTGGTCCAGTGGCTTGGGGGCGAGGCGAGTGA 804
| | | | |
Db 7569 GAGTGGGCGGTGTGTGCTATGCCGTATCCAGCACTTGGGAGGCCAAAGCAGGTGA 7628
| | | | |
QY 805 TCACCTTAAGTCAGAGATTGAGACCAAGCCTGCCCAACATGTTGAAACGTTGCTACT 864
| | | | |
Db 7629 TCAGTTGGGTCAAGAGCTTGAGACCAAGCTGGCCCAACATGTTGAAACGTTGCTACT 7688
| | | | |
QY 865 AAAAAATCAAAAAATTAGACAGGCGCTGTGGGACACATCTGTAATTCAGCTACTCAGAG 924
| | | | |
Db 7689 AAAAAATCAAAAAATTAGACAGAGTGTGTGGGCGCACTTGTAAATCCAGCTACTCGGAG 7748
| | | | |
QY 925 GCTTAACAGAGAAATTCCTTGAACCTGAGAGGAGAGGTTGCACTGA-----G 973
| | | | |
Db 7749 GCTAGGCAAGAAATGCTTGAACCAAGAGGTGAGAGTTGCACTGAGCCGAGATCATG 7808
| | | | |
QY 974 CCATTGCACTCCAGCTGGGCAACAGTGAAGCTTGTCTCAAAAAAAAAAAAAA 1032
| | | | |
Db 7809 CCGGTGACTCCAGCCTGGGTGACAGAGCGAGACTTGTCTCAAAAAAAAAAAAAA 7867
| | | | |

RESULT 12
US-09-949-016-14222
; Sequence 14222, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14222
; LENGTH: 88736
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14222

Query Match 19.1%; Score 198.2; DB 4; Length 88736;
Best Local Similarity 81.9%; Pred. No. 2.4e-38;
Matches 245; Conservative 0; Mismatches 43; Indels 11; Gaps 1;
QY 745 GAGTGGGCGCATAGCTGATGCTGTGGTCCAGTGGCTTGGGGGCGAGGCGAGTGA 804
| | | | |
Db 7815 GAGTGGGCGGTGTGTGCTATGCCGTATCCAGCACTTGGGAGGCCAAAGCAGGTGA 7874
| | | | |
QY 805 TCACCTTAAGTCAGAGATTGAGACCAAGCCTGCCCAACATGTTGAAACGTTGCTACT 864
| | | | |
Db 7875 TCAGTTAGGTCAAGAGCTTGAACCAAGCTGGCCCAACATGTTGAAACGTTGCTACT 7934
| | | | |
QY 865 AAAAAATCAAAAAATTGACAGGCGTGTGGCAACATCTGTATTCAGCTACTCAGAG 924
| | | | |
Db 7935 AAAAAATCAAAAAATTGACAGAGTGTGGCGGCACTGTATCCAGCTACTCGGAG 7994
| | | | |
QY 925 GCTTAACAGAGAAATTCCTTGAACCTGGGAGGAGAGTTGCACTGA-----G 973
| | | | |
Db 7995 GCTAGGCAAGAAATGCTTGAACCAAGAGTGTGAGAGTTGCACTGAGCCGAGATCATG 8054
| | | | |
QY 974 CCATTGCACTCCAGCTGGGCAACAGTGAAGCTTGTCTCAAAAAAAAAAAAAA 1032
| | | | |
Db 8055 CCGGTGACTCCAGCCTGGGTGACAGAGCGAGACTTGTCTCAAAAAAAAAAAAAA 8113
| | | | |

RESULT 13
US-09-949-016-45906/C
; Sequence 45906, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45906
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-45906

Query Match 19.0%; Score 196.8; DB 4; Length 601;
Best Local Similarity 79.8%; Pred. No. 9.4e-39;
Matches 249; Conservative 0; Mismatches 52; Indels 11; Gaps 1;


```
QY 733 TTTTACGACGAGGCTGGGCGCATAGCTGATGCTGNGGATCCGAGNCTTTGCGGGGCC 792
DB 596 TTATATTTTAAAGCCCGGCGATGCTGCTCAGCGCTTATATCCAGACATCTTGGAGGCT 537
QY 793 GAGGACAGTGCATCTTAAGGTCAGAGTTTGAACAGCCTTCCCAACATGCTGAAC 852
DB 536 GAGGTAGCGGATCACTGAGGTGAGAGTTGAGACAGCTTGGCAACATGCTGAAT 477
QY 853 GTTGTCTCTAATAAATAAATAAATTAGACAGCGCTGGTGGCACATCTGTAATCCA 912
DB 476 CCGTCCCTCTAATAAATAAATAAATTAGCAGATGCTGGCATGCTGTAATCCA 417
QY 913 GCTACTGAGAGGCTAACAAGAAATTCCTTGAACCTGGGAGGACAGAGTTGCACTGA 972
DB 416 GCTACTGAGAGGCTGAGGAGGAAATGCTTGAACCTGGGAGGACAGAGTTGCACTGA 357
QY 973 GC-----CATTCGACTCCAGCCTGGGCAACACAGTGAACCTTGTCTCAAAA 1021
DB 356 GCGGAGATCGCATCACTGCACTCCAGCCTAGTGAACAGAGTGAACCTTGTCTCAAAA 297
QY 1022 AAAAAAAAAA 1033
DB 296 AAAAAAAAAA 285

RESULT 14
US-09-949-016-13052
; Sequence 13052, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13052
; LENGTH: 20441
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)..(20441)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13052

Query Match 19.0%; Score 196.8; DB 4; Length 20441;
Best Local Similarity 79.8%; Pred. No. 3.2e-38;
Matches 249; Conservative 0; Mismatches 52; Indels 11; Gaps 1;
```

```
DB 11864 GCTACTCAGAGGCTGAGGAGAGAAATTCCTTGAACCTGGGAGGACAGAGTTGCACTGA 11923
QY 973 GC-----CATTCGACTCCAGCCTGGGCAACACAGTGAACCTTGTCTCAAAA 1021
DB 11924 GCGGAGATCGCATCACTGCACTCCAGCCTAGTGAACAGAGTGAACCTTGTCTCAAAA 11983
QY 1022 AAAAAAAAAA 1033
DB 11984 AAAAAAAAAA 11995

RESULT 15
US-09-949-016-45905/c
; Sequence 45905, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 45905
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-45905

Query Match 18.9%; Score 196.4; DB 4; Length 601;
Best Local Similarity 79.5%; Pred. No. 1.2e-38;
Matches 248; Conservative 1; Mismatches 52; Indels 11; Gaps 1;
```

```
QY 733 TTTTACGACGAGGCTGGGCGCATAGCTGATGCTGTGCTCCAGTCTTTGCGGGGCC 792
DB 579 TTATATTTTAAAGCCCGGCGATGCTGCTCAGCGCTTATATCCAGACATCTTGGAGGCT 520
QY 793 GAGGACAGTGCATCTTAAGGTCAGAGTTTGAACAGCCTTCCCAACATGCTGAAC 852
DB 519 GAGGTAGCGGATCACTGAGGTGAGAGTTGAGACAGCTTGGCAACATGCTGAAT 460
QY 853 GTTGTCTCTAATAAATAAATAAATTAGACAGCGCTGGTGGCACATCTGTAATCCA 912
DB 459 CCGTCCCTCTAATAAATAAATAAATTAGCAGATGCTGGCATGCTGTAATCCA 400
QY 913 GCTACTGAGAGGCTAACAAGAAATTCCTTGAACCTGGGAGGACAGAGTTGCACTGA 972
DB 399 GCTACTGAGAGGCTGAGGAGGAAATTCCTTGAACCTGGGAGGACAGAGTTGCACTGA 340
QY 973 GC-----CATTCGACTCCAGCCTGGGCAACACAGTGAACCTTGTCTCAAAA 1021
DB 339 GCGGAGATCGCATCACTGCACTCCAGCCTAGTGAACAGAGTGAACCTTGTCTCAAAA 280
QY 1022 AAAAAAAAAA 1033
DB 279 AAAAAAAAAA 268

Search completed: March 6, 2005, 06:42:04
Job time : 224 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 6, 2005, 06:38:25 ; Search time 121 Seconds
(without alignments)
156.622 Million cell updates/sec

Title: US-10-664-358-549

Perfect score: 260

Sequence: 1 MWLGMIFSMCGMLKLKMC.....ISFANSRSSBDTKQWMSSEF 49

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1980s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	49	3 AAY87141	AAY87141 Human sec
2	260	100.0	49	4 AAB06118	AAB06118 Human gen
3	260	100.0	49	5 ABG33940	ABG33940 Human sec
4	260	100.0	49	6 ABR47655	ABR47655 Human sec
5	260	100.0	49	6 ABR00022	ABR00022 Human gen
6	260	100.0	49	7 ADB91468	ADB91468 Human sec
7	260	100.0	49	7 ADC74024	ADC74024 Human sec
8	260	100.0	50	2 AAW67873	AAW67873 Human sec
9	260	100.0	70	3 AAY76058	AAY76058 Human skt
10	260	100.0	70	4 AAB55997	AAB55997 Skin cell
11	260	100.0	70	4 AAE06054	AAE06054 Human gen
12	260	100.0	70	5 ABG33876	ABG33876 Human sec
13	260	100.0	70	5 ABB72197	ABB72197 Human pro
14	260	100.0	71	3 AAY87077	AAY87077 Human sec
15	260	100.0	106	2 AAY59659	AAy59659 Secreted
16	260	100.0	106	3 AAY94891	AAy94891 Human pro
17	260	100.0	106	5 AAG04034	AAG04034 Human sec
18	260	100.0	106	5 ABB89692	ABB89692 Human pol
19	260	100.0	106	7 ADM77809	ADM77809 Human pro
20	260	100.0	106	8 ADP19490	ADP19490 Human sec
21	260	100.0	106	8 ADQ95956	ADQ95956 T cell ac
22	260	100.0	112	4 AAB43968	AAB43968 Human can
23	260	100.0	132	4 ABB11978	ABB11978 Human sec
24	217	83.5	108	4 ABB65197	ABB65197 Drosophila
25	158	60.8	30	2 AAW67999	AAW67999 Fragment

26	110	42.3	99	4 AAU32155	AAU32155 Novel hum
27	97	37.3	20	2 AAW67997	AAW67997 Fragment
28	73.5	28.3	107	3 AAG09358	AAG09358 Arabidops
29	61	23.5	308	3 AAB43188	AAB43188 Human ORF
30	61	23.5	360	7 ADB65455	ADB65455 Human ORF
31	61	23.5	406	5 ADI16911	ADI16911 NOX prot
32	61	23.5	415	8 ADI16586	ADI16586 Human NOV
33	61	23.5	415	8 ADN42240	ADN42240 Human nov
34	61	23.5	416	8 ADI16588	ADI16588 Human NOV
35	61	23.5	416	8 ADN42242	ADN42242 Human nov
36	59.5	22.9	57	3 AAG09360	AAG09360 Arabidops
37	59.5	22.9	1970	4 ABB64827	ABB64827 Drosophila
38	59	22.7	432	8 ADN18428	ADN18428 Bacterioph
39	58.5	22.5	332	2 AAY30160	AAY30160 Human dor
40	58.5	22.5	733	4 ABB24605	ABB24605 Novel hum
41	58	22.3	234	4 ABB21806	ABB21806 Novel hum
42	58	22.3	391	7 ADB64644	ADB64644 Human pro
43	57.5	22.1	193	4 AAG71944	AAG71944 Human Olf
44	57	21.9	358	5 ABB65847	ABB65847 Blifidobac
45	57	21.9	472	4 ABB65583	ABB65583 Drosophila

ALIGNMENTS

RESULT 1
ID AAY87141 standard; protein; 49 AA.
AC AAY87141;
DT 09-MAY-2000 (first entry)
DE Human secreted protein sequence SEQ ID NO:180.
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;
XX tumour; neurodegenerative disorder; developmental abnormality; allergy;
XX focal deficiency; blood disorder; immune system disorder; arthritis;
XX autoimmune disease; hepatic disease; renal disease; inflammation;
XX Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
XX infection; AIDS; spinal cord injury; transplant rejection; diabetes;
XX asthma; sepsis; acne; psoriasis; cardiovascular disorder;
XX reproductive disorder; gastrointestinal disorder; respiratory disorder;
XX metabolic disorder; food additive; preservative.
XX Homo sapiens.
XX OS
XX PN WO200004140-A1.
XX 27-JAN-2000.
XX 14-JUL-1999; 99WO-US015849.
XX PF 15-JUL-1998; 98US-0092921P.
XX PR 15-JUL-1998; 98US-0092922P.
XX PR 15-JUL-1998; 98US-0092956P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
XX Lafleur DW, Edner R, Olsen HS, Brewer LA, Florence KA, Young PE;
XX Mucenski M, Endress GA, Soppet DR;
XX WPI; 2000-161128/14.
XX DR N-PSDB; AA298094.
XX PT New isolated human genes, useful for diagnosis and treatment of, e.g.
XX PT cancers, neurological or blood disorders.
XX PS Claim 11; Page 429; 494pp; English.
XX CC The polynucleotide sequences given in AA298017 to AA298108 encode the

CC human secreted proteins given in AAY87064 to AAY87223. Human secreted
CC protein can have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: cytostatic; and
CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
CC antiallergic. The polynucleotides and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new
CC polynucleotides. Human secreted proteins and their polynucleotides can
CC be used for developing products for the diagnosis or treatment of cancer,
CC tumours, neurodegenerative disorders, developmental abnormalities and
CC foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
CC Alzheimer's disease, behavioural disorders, schizophrenia, osteoporosis,
CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
CC reproductive disorders, gastrointestinal disorders, respiratory disorders
CC and metabolic disorders. The proteins or polynucleotides can also be used
CC as food additives or preservatives. The proteins are also useful for
CC identifying their binding partners. AAZ86008 to AAZ98016 and AAY87063 are
CC sequence used in the exemplification of the present invention

CC Sequence 49 AA;

Query Match 100.0%; Score 260; DB 3; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30; Indels 0; Gaps 0;
Matches 49; Conservative 0; Mismatches 0;

Qy 1 MNLIGMIFSMCGIMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49
Db 1 MNLIGMIFSMCGIMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49

RESULT 2
AAE06118 standard; protein; 49 AA.

XX AAE06118;

XX 24-SEP-2001 (first entry)

DE Human gene 14 encoded secreted protein HAU183, SEQ ID NO:180.

KM Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
KM foetal abnormality; developmental abnormality; haematopoietic disorder;
KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KM Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
KM psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
KM inflammation; neurological disorder; Alzheimer's disease; food additive;
KM angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
KM pregnancy-related disorder; endocrine disorder; infection; wound healing;
KM cell culture; chemotaxis; vulnery; binding partner identification;
KM gene therapy; chromosome 19.

XX Homo sapiens.

OS Key Location/Qualifiers
FH Peptide 1..37
FT /label= signal_peptide
FT Protein 38..49
FT /label= Mature_human_secreted_protein

XX MO200151504-A1.

XX 19-JUL-2001.

PD 12-JAN-2001; 2001WO-US000911.

PF 13-JAN-2000; 2000US-00482273.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PA
XX XX

PI Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
PI Lefler DM, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
PI Endress GA, Muscenski M, Edner R;
XX WPI; 2001-425865/45.
DR N-PSDB; AAD11707.

XX Isolated nucleic acid molecule encoding a human secreted protein is used
PT in preventing, treating or ameliorating a medical condition.

XX Claim 11; Page 796; 864pp; English.

XX AAD11630-AAD11721 represent CDNA's corresponding to 71 human secreted
CC protein genes, AAE06041-AAE06132 represent the proteins they encode.
CC AAE06133-AAE06205 represent human secreted protein fragments. The
CC secreted proteins and their genes are useful for preventing, treating or
CC ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 71 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays (e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA)). The
CC present sequence represents a human secreted protein of the invention

XX Sequence 49 AA;

Query Match 100.0%; Score 260; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30; Indels 0; Gaps 0;
Matches 49; Conservative 0; Mismatches 0;

Qy 1 MNLIGMIFSMCGIMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49
Db 1 MNLIGMIFSMCGIMLKIKWCAMVAVYCSFISFANSRSSEDTKOMSSFM 49

RESULT 3
ABG33940 standard; protein; 49 AA.

XX ABG33940;

XX 15-JUL-2002 (first entry)

DE Human secreted protein encoded by gene 14 #2.

XX Human; secreted protein; gene therapy; immunosuppressive; antiarthritic;
KM antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KM cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KM fungicide; ophthalmological; autoimmune disease; neoplasm;
KM rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KM cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KM angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KM ocular disorder; corneal infection; wound healing; skin aging;
KM epithelial cell proliferation; food additive.

XX Homo sapiens.

OS

XX WO200226931-A2.
PN
XX
XX 04-APR-2002.
PD
XX
XX 24-SEP-2001; 2001WO-US029871.
PP
XX 25-SEP-2000; 2000US-0234925P.
PR
XX 12-JAN-2001; 2001WO-US000911.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y,
PI Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR,
PI Endress GA, Muceneki M, Ebner R;
XX WPI; 2002-362489/39.
DR N-PSDB; ABK69803.
XX
XX Novel 71 isolated secreted polypeptides and polynucleotides encoding the
PT polypeptides, useful for treating Huntington's disease, sepsis,
PT meningitis, thrombocytopenia, hemolytic anemia, rheumatoid arthritis,
PT asthma.
XX
XX Claim 11; Page 1273; 1478pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (or its
XX fragment, homologue complement or allelic variant) encoding a human
CC secreted protein (and its fragment, domain, epitope, variant, secreted
CC form and species variant). Also included are a recombinant vector
CC comprising the nucleic acid, a recombinant host cell comprising the
CC vector, an antibody against the secreted protein, a recombinant host cell
CC that expresses the secreted protein and a method of identifying a binding
CC partner of the secreted protein. The nucleic acid and protein are used to
CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep for
CC example autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,
CC angiodementia, nervous system disorders e.g. Alzheimer's disease,
CC infections caused by bacteria, viruses and fungi and ocular disorders
CC e.g. corneal infection. Many other diseases and disorders are listed in
CC the specification. The polypeptides can also be used to aid wound healing
CC an epithelial cell proliferation, to prevent skin aging due to sunburn, of
CC to maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The present sequence
CC represents a novel human secreted protein of the invention
XX
XX
XX Sequence 49 AA;
SQ
Query Match 100.0%; Score 260; DB 5; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLIGMIFSMCGMLTKIKWCANVAVYCSFISFANSRSSEDTKOMSSFM 49
Db 1 MNLIGMIFSMCGMLTKIKWCANVAVYCSFISFANSRSSEDTKOMSSFM 49

RESULT 4
ABR47655
ID ABR47655 standard; protein; 49 AA.
XX
XX ABR47655;
AC
XX
XX 12-JUN-2003 (first entry)
DT
XX
XX Human secreted protein, SEQ ID 546.
DE
XX
XX Cardiac; antiarrhythmic; antiarteriosclerotic; vasotrophic; cyostatic;
KW vulnerable; antiinflammatory; nootropic; neuroprotective;

KW antiparkinsonian; gene therapy; human; cardiovascular disorder.
XX
XX Homo sapiens.
OS
XX
XX WO200295010-A2.
FN
XX
XX 28-NOV-2002.
PD
XX
XX 19-MAR-2002; 2002WO-US009785.
PP
XX
XX 21-MAR-2001; 2001US-0277340P.
PR
XX 19-JUL-2001; 2001US-0306171P.
PR
XX 13-NOV-2001; 2001US-0331287P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Ruben SM;
PI
XX WPI; 2003-129429/12.
DR
XX
XX Novel human secreted proteins, useful for detecting, preventing,
PT diagnosing, prognosticating, treating and/or ameliorating cardiovascular
PT disorders such as arrhythmia.
XX
XX Claim 13; SEQ ID NO 546; 1881pp; English.
PS
XX

XX The present invention relates to novel human secreted proteins (ABR47653-
CC ABR48145) and their coding sequences (ACCS0344-ACCS05056). The proteins
CC and their coding sequences are useful for the preparation of a diagnostic
CC or pharmaceutical composition for diagnosing or treating a cardiovascular
CC disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
CC arteriosclerosis and myocardial ischemia), neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, renal disorders,
CC proliferative disorders and/or cancerous diseases and conditions, for
CC wound healing and epithelial cell proliferation, to treat inflammation or
CC infection, for treating thrombosis and arteriosclerosis, for treating or
CC preventing neural damage which occurs in neuronal disorders or
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
CC disease, to enhance bone and periodontal regeneration and aid in tissue
CC transplants or bone grafts, to prevent skin aging or hair loss, to
CC stimulate growth and differentiation of haematopoietic cells and bone
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism. Note: The sequence data for this patent was published in
CC electronic format and is available from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
XX Sequence 49 AA;
SQ

Qy 1 MNLIGMIFSMCGMLTKIKWCANVAVYCSFISFANSRSSEDTKOMSSFM 49
Db 1 MNLIGMIFSMCGMLTKIKWCANVAVYCSFISFANSRSSEDTKOMSSFM 49

RESULT 5
ABR00022
ID ABR00022 standard; protein; 49 AA.
XX
XX ABR00022;
AC
XX
XX 03-APR-2003 (first entry)
DT
XX
XX Human gene 12 encoded secreted protein HAVUA183, SEQ ID NO:311.
DE
XX
XX Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;

KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic; gene therapy;
KW antiinflammatory; immunosuppressive; vulnery; chromosome 19.
OS Homo sapiens.
XX
XX WO200276488-A1.
PN
PD 03-OCT-2002.
XX
XX 19-MAR-2002; 2002WO-US008276.
PF
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-029900/02.
DR N-PSDB; ABZ71201.
XX
XX
PT New human secreted proteins and nucleic acids, useful for detecting,
PT preventing, diagnosing, prognosticating, treating and/or ameliorating
PT e.g. gastrointestinal diseases and disorders, or cancers.
XX
XX
XX Claim 13; Page 942; 1216pp; English.
XX
XX ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
XX protein genes, and ABP00011-ABP00299 represent the proteins they encode.
XX ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
XX invention also encompasses antibodies specific for the secreted proteins, the
XX use of the secreted proteins in drug screening, and recombinant
XX vectors and host cells comprising a nucleic acid of the invention. The
XX secreted proteins, nucleic acids encoding them, antibodies or antibody
XX fragments specific for the secreted proteins, and modulators of protein
XX activity are useful for diagnosing, treating, ameliorating or preventing
XX digestive disorders. Such conditions include disorders of the mouth,
XX oesophagus, stomach, small intestine, large intestine, liver, biliary
XX tract and pancreas, and include cancers of these organs and tissues. The
XX secreted proteins and their nucleic acids may also be used in the
XX treatment of immune disorders, inflammation, infection,
XX hyperproliferative disorders, and to promote wound healing. Nucleic acids
XX of the invention may be used for chromosome identification, chromosome
XX mapping, in gene therapy, for identifying individuals from minute
XX biological samples, as hybridisation probes, and as molecular weight
XX markers. The present sequence represents a human secreted protein of the
XX invention
SQ Sequence 49 AA;
Query Match 100.0%; Score 260; DB 6; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NMLIGMIFSMGMLKLKRCAMVAAYCSFISFANSRSSEDTKQMMSSFM 49
DB 1 NMLIGMIFSMGMLKLKRCAMVAAYCSFISFANSRSSEDTKQMMSSFM 49
RESULT 6
ADB91468
ID ADB91468 standard; protein; 49 AA.
AC ADB91468;
XX
XX 04-DEC-2003 (first entry)
DT
XX
XX Human secreted protein #SEQ ID 414.
DE
XX
KW Secreted protein; gene therapy; antidiabetic; diabetes; human.

XX
OS Homo sapiens.
XX
XX WO2003004622-A2.
PN
XX
XX 16-JAN-2003.
XX
XX 19-MAR-2002; 2002WO-US008124.
PF
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-229407/22.
DR
XX
XX
PT Nucleic acid encoding a human secreted protein is useful in diagnosing or
PT treating diabetes or conditions related to diabetes.
XX
XX
XX Claim 3; SEQ ID NO 414; 1537pp; English.
XX
XX The invention relates to isolated nucleic acid molecules ADB91065-
XX ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
XX ADB91834. Also disclosed is a recombinant vector comprising a
XX polynucleotide of the invention, and a recombinant host cell comprising
XX the recombinant vector. The polypeptide of the invention is useful in
XX identifying a binding partner by contacting the polypeptide with a
XX binding partner, and determining whether the binding partner increases or
XX decreases activity of the polypeptide. The polypeptide, polynucleotide,
XX antibody or its fragment, agonist or antagonist are useful for preparing
XX a pharmaceutical composition for diagnosing or treating diabetes or
XX conditions related to diabetes. The present sequence is that of the human
XX immunoglobulin Fc portion used to generate fusion proteins, increasing
XX the stability of the fused protein as compared to the secreted protein
XX only. Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at http://wipo.int/pub/published_pct_sequences.
SQ Sequence 49 AA;
Query Match 100.0%; Score 260; DB 7; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.9e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NMLIGMIFSMGMLKLKRCAMVAAYCSFISFANSRSSEDTKQMMSSFM 49
DB 1 NMLIGMIFSMGMLKLKRCAMVAAYCSFISFANSRSSEDTKQMMSSFM 49
RESULT 7
ADCT74024
ID ADCT74024 standard; protein; 49 AA.
AC ADCT74024;
XX
XX 01-JAN-2004 (first entry)
DT
XX
XX Human secreted protein - SEQ ID 657.
DE
XX
XX antianemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
KW antidiabetic; immunosuppressive; dermatological; nephrotoxic;
KW antiparkinsonian; neuroprotective; nootropic; antibacterial; virocidic;
KW fungicide; antiparasitic; antiarteriosclerotic; vulnery; cytostatic;
KW haemopoietic; haematologic; anaemia; autoimmune disorder;
KW rheumatoid arthritis; inflammation; Grave's disease; diabetes;
KW systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
KW Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
KW cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
XX human.

OS Homo sapiens.
XX
PN WO2003038663-A2.
XX
PD 08-MAY-2003.
XX
PF 19-MAR-2002; 2002WO-US008277.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
XX WPI: 2003-430516/40.
DR N-PSDB; ADC73409.
XX
XX New human secreted polypeptide for diagnosing, preventing or treating
PT hematopoietic or hematologic disorders (e.g. anemia), autoimmune
PT disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
PT atherosclerosis).
XX
XX Claim 16; SEQ ID NO 657; 2272pp; English.
XX
XX The invention relates to a novel human secreted polypeptide comprising a
CC defined sequence given in the specification. The polypeptide, nucleic
CC acid molecule, antibody, agonist or antagonist of the invention may be
CC useful for preparing a composition for diagnosing or treating a
CC hematopoietic or hematologic disorder such as anemia, autoimmune
CC disorders such as rheumatoid arthritis, inflammation, Grave's disease,
CC diabetes, systemic lupus erythematosus or glomerulonephritis,
CC neurodegenerative disorders including Parkinson's disease and Alzheimer's
CC disease, wounds and hyperproliferative disorders including
CC parasitic infections. The polypeptide may also be used during gene
CC therapy procedures and for identifying a binding partner by contacting
CC the polypeptide with a binding partner and determining whether the
CC binding partner increases or decreases the activity of the polypeptide.
CC The current sequence is that of the human secreted protein of the
CC invention.
XX
XX Sequence 49 AA;
SQ
Query Match 100.0%; Score 260; DB 7; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.3e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNLIGMIFSMCGMLKXKMCAMWAVVCSFISFANSRSSEDTKOMSSFM 49
DB 1 MNLIGMIFSMCGMLKXKMCAMWAVVCSFISFANSRSSEDTKOMSSFM 49
RESULT 8
AAW67873 ID AAW67873 standard; protein; 50 AA.
XX
XX AAW67873;
XX
DT 25-MAR-1999 (first entry)
XX
XX Human secreted protein.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KM diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KM developmental abnormality; foetal deficiency; blood; allergy; renal;
KM immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KM inflammation; ischaemic shock; Alzheimer's disease; resveratrol; AIDS;
KM cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KM osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KM endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX

OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 50
FT /Label= unknown
XX
XX WO9842738-A1.
XX
XX 01-OCT-1998.
XX
XX 19-MAR-1998; 98WO-US005311.
XX
XX 21-MAR-1997; 97US-0041276P.
XX 21-MAR-1997; 97US-0041277P.
XX 21-MAR-1997; 97US-0041281P.
XX 21-MAR-1997; 97US-0042344P.
XX 30-MAY-1997; 97US-0048069P.
XX 30-MAY-1997; 97US-0048094P.
XX 30-MAY-1997; 97US-0048095P.
XX 30-MAY-1997; 97US-0048096P.
XX 30-MAY-1997; 97US-0048099P.
XX 30-MAY-1997; 97US-0048131P.
XX 30-MAY-1997; 97US-0048135P.
XX 30-MAY-1997; 97US-0048154P.
XX 30-MAY-1997; 97US-0048160P.
XX 30-MAY-1997; 97US-0048166P.
XX 30-MAY-1997; 97US-0048187P.
XX 30-MAY-1997; 97US-0048188P.
XX 30-MAY-1997; 97US-0048351P.
XX 30-MAY-1997; 97US-0048352P.
XX 30-MAY-1997; 97US-0048353P.
XX 30-MAY-1997; 97US-0048355P.
XX 30-MAY-1997; 97US-0050937P.
XX 05-AUG-1997; 97US-0054804P.
XX 19-AUG-1997; 97US-0056370P.
XX 02-OCT-1997; 97US-0060862P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Duan R, Hu J;
PI Florence KA, Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y;
PI Lafleur DW, Ni J;
XX
XX WPI: 1999-070066/06.
DR N-PSDB; AAX00677.
XX
XX New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
XX Claim 11; Page 308; 385pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AAX00602) for increasing the stability of the fused protein
CC as compared to the human protein only. The invention relates to 87 novel
CC genes and their fragments (nucleic acid sequences: AAX00611-X00724; amino
CC acid sequences AAW67807-W68004) which are useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
CC pathological conditions can be diagnosed by determining the amount of the
CC new polypeptides in a sample or by determining the presence of mutations
CC in the new polynucleotides. Specific uses are described for each of the
CC 87 polynucleotides, based on which tissues they are most highly expressed
CC in (see AAX00611 for described uses)
XX
XX Sequence 50 AA;
SQ
Query Match 100.0%; Score 260; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 3e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNLIGMIFSMCGMLKXKMCAMWAVVCSFISFANSRSSEDTKOMSSFM 49

Db 1 MNLIGMIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSDTKOMSSFM 49

RESULT 9

AA76058
ID AAY76058 standard; protein; 70 AA.

AC AAY76058;

DT 27-MAR-2000 (first entry)

DE Human skin cell protein, SEQ ID NO:313.

XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW secreted; transmembrane; inflammation; cancer; neurological disease;
KW angiogenesis; tumor vasculatization; growth disorder;
KW developmental disorder; skin wound; hair follicle disorder;
KW anti-inflammatory; cytostatic; neuroprotective; vulnery.

OS Homo sapiens.

PN WO955865-A1.

PD 04-NOV-1999.

PF 29-APR-1999; 99WO-NZ000051.

PR 29-APR-1998; 98US-00069726.

PR 09-NOV-1998; 98US-00188930.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;

DR WPI; 2000-072177/06.

PT Novel polynucleotides useful for the treatment of various conditions
including wounds and cancer.

PS Claim 4; Page 185; 235pp; English.

XX The invention relates to novel nucleic acid sequences derived from rat
XX dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC cells. Polypeptides of the invention may be used to treat inflammation,
CC cancer and neurological diseases. The proteins may be used to stimulate
CC the growth and motility of keratinocytes, to inhibit the growth of cancer
CC cells, to modulate angiogenesis and tumor vasculatization, to modulate
CC skin inflammation, to modulate epithelial cell growth and to inhibit
CC binding of HIV-1 to leukocytes. The invention may also be used to treat
CC growth and developmental defects, skin wounds and hair follicle
CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded by
CC cDNA sequences derived from several mouse, rat or human skin cell types.
CC Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and AAY76119
CC are proteins with an N-terminal signal sequence, indicating that they are
CC secreted. Sequences AAY75986-Y75989, AAY76061-Y76071, AAY76106-Y76109 and
CC AAY76121-Y76122 are proteins with one or more putative transmembrane
CC domains

XX Sequence 70 AA;

XX Query Match 100.0%; Score 260; DB 3; Length 70;

XX Best Local Similarity 100.0%; Pred. No. 4.3e-30;

XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSDTKOMSSFM 49

Db 1 MNLIGMIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSDTKOMSSFM 49

RESULT 10

AAB55997

ID AAB55997 standard; protein; 70 AA.

XX AAB55997;

XX 08-MAR-2001 (first entry)

XX Skin cell protein, SEQ ID NO: 313.

XX Human; skin cell; cytostatic; antiinflammatory; anti-HIV; nootropic;
KW neuroprotective; vulnery; immunomodulatory; vaccine;
KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW inflammation; neurological disease.

XX Homo sapiens.

XX WO200069884-A2.

XX 23-NOV-2000.

XX 15-MAY-2000; 2000WO-NZ000075.

XX 14-MAY-1999; 99US-00312283.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PI Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD;

PI Murison JG;

XX WPI; 2001-007495/01.

XX N-FSD8; AAC99696.

XX New isolated polynucleotide used in the identification of genetic
PT disorders and encoding polypeptides used for treating inflammatory
PT disease, cancer and neurological diseases.

PS Claim 4; Page 251; 352pp; English.

XX The present sequence is a polypeptide which is expressed in mammalian
CC skin cells. The polypeptide is useful for stimulating keratinocyte growth
CC and motility, inhibiting the growth of cancer cells, modulating
CC angiogenesis, inhibiting angiogenesis and vasculatization of tumours,
CC modulating skin inflammation, stimulating the growth of epithelial cells,
CC inhibiting the binding of human immunodeficiency virus (HIV)-1 to
CC leukocytes, and treating inflammatory disease, cancer and neurological
CC diseases. The polynucleotide can be used as a marker, in the
CC identification of genetic disorders, and for the design of
CC oligonucleotides for examining expression patterns

XX Sequence 70 AA;

XX Query Match 100.0%; Score 260; DB 4; Length 70;

XX Best Local Similarity 100.0%; Pred. No. 4.3e-30;

XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSDTKOMSSFM 49

Db 1 MNLIGMIFSMCGIMLKLKWCAMVAVYCSFISFANSRSSSDTKOMSSFM 49

RESULT 11

AAE06054

XX AAE06054 standard; protein; 70 AA.

XX AAE06054;

XX 24-SEP-2001 (first entry)

XX Human gene 14 encoded secreted protein HHEPG41, SEQ ID NO:116.
DE Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
KW foetal abnormality; developmental abnormality; hematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;

KW		psoctiaies; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
KW		inflammation; neurological disorder; Alzheimer's disease; food additive;
KM		angiotonic disorder; kidney disorder; gastrointestinal disorder; allergy;
KM		pregnancy-related disorder; endocrine disorder; infection; wound healing;
KM		cell culture; chemotaxis; vulneryary; binding partner identification;
KX		gene therapy; chromosome 19.
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..35
FT	Protein	/label= signal_peptide 36..70
FT		/label= Mature_human_secreted_protein
PN	MO200151504-AI.	
PD	19-JUL-2001.	
XX		
PE	12-JAN-2001; 2001WO-US000911.	
XX		
PR	13-JAN-2000; 2000US-00482273.	
XX		
PA	(HUMA-) HUMAN GENOME SCT INC.	
XX		
P1	Rubben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y,	
P1	Lafleur DM, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR,	
P1	Endress GA, Musemekl M, Ebner R,	
XX		
DR	WPI, 2001-425865/45.	
XX	N-PDSB; AAD11643.	
PT		
PT	Isolated nucleic acid molecule encoding a human secreted protein is used	
PS	in preventing treating or ameliorating a medical condition.	
PS	Claim 11; Page 747; 864pp; English.	
XX		
CC	AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted	
CC	protein genes, and AAEO6041-AAEO6132 represent the proteins they encode.	
CC	AAE06133-AAE06205 represent human secreted protein fragments. The	
CC	secreted proteins and their genes are useful for preventing, treating or	
CC	ameliorating medical conditions, e.g., by protein or gene therapy.	
CC	Pathological conditions can be diagnosed by determining the amount of the	
CC	new protein in a sample or by determining the presence of mutations in	
CC	the new genes. Specific uses are described for each of the 71 genes,	
CC	based on the tissues in which they are most highly expressed, and include	
CC	developing products for the diagnosis or treatment of proliferative	
CC	disorders, cancer, tumours, foetal and developmental abnormalities,	
CC	haematopoietic disorders, diseases of the immune system, AIDS, autoimmune	
CC	diseases (e.g., rheumatoid arthritis), inflammation, allergies,	
CC	neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),	
CC	cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,	
CC	psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,	
CC	angiogenic disorders, kidney disorders, gastrointestinal disorders,	
CC	pregnancy-related disorders, endocrine disorders, and infections. The	
CC	proteins can also be used to aid wound healing and epithelial cell	
CC	proliferation, to prevent skin aging due to sunburn, to maintain organs	
CC	before transplantation, for supporting cell culture of primary tissues,	
CC	to regenerate tissues, to identify their cognate ligands or binding	
CC	partners, and in chemotaxis, and can be used as a food additive or	
CC	preservative to modify storage properties. Antibodies specific for a	
CC	protein of the invention can be used in alleviating symptoms associated	
CC	with the disorders mentioned above, and in diagnostic immunoassays e.g.,	
CC	radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The	
CC	present sequence represents a human secreted protein of the invention	
XX		
SQ	Sequence 70 AA:	
XX		
Query Match	100.0%; Score 260; DB 4; Length 70;	
Match Local Similarity	100.0%; Pred. No. 4.3e-30;	
Matches 49; Conservative	0; Mismatches 0; Indels 0; Gaps 0	

Db 1 |||||

RESULT 12
ABG33876
ID ABG33876 standard; protein; 70 AA.
XX
XX ABG33876;
DT 15-JUL-2002 (first entry)
DE Human secreted protein encoded by gene 14 #1.
KW Human; secreted protein; gene therapy; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cyostatic; cardiac; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; vitucide;
KW fungicide; ophthalmological; autoimmune disease; neoplasm;
KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW ocular disorder; corneal infection; wound healing; skin aging;
KW epithelial cell proliferation; food additive.
XX
XX Homo sapiens.
OS
XX WO200226931-A2.
PN
XX
XX 04-APR-2002.
PD
XX
XX 24-SEP-2001; 2001WO-US029871.
PE
XX 25-SEP-2000; 2000US-0234925P.
PR 12-JAN-2001; 2001WO-US000911.
RX
XX (HUMA-) HUMAN GENOME SCT INC.
PA
PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y,
PI Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR,
PI Endress GA, Mucenski M, Ebner R;
PT WPI; 2002-362489/39.
DR N-PSDB; ABE69739.
XX
XX Novel 71 isolated secreted polypeptides and polymucleotides encoding the
PT polypeptides, useful for treating Huntington's disease, sepsis,
PT meningitis, thrombocytopenia, hemolytic anemia, rheumatoid arthritis,
PT asthma.
XX
XX Claim 11; Page 1228; 1478pp; English.

The invention relates to an isolated nucleic acid molecule (or its fragment, homologue complement or allelic variant) encoding a human secreted protein (and its fragment, domain, epitope, variant, secreted form and species variant). Also included are a recombinant vector comprising the nucleic acid, a recombinant host cell comprising the vector, an antibody against the secreted protein, a recombinant host cell that expresses the secreted protein. The nucleic acid and protein are used to partner of the secreted protein. The nucleic acid and protein are used to prevent, diagnose, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep for example autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angioneuosis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. Many other diseases and disorders are listed in the specification. The polypeptides can also be used to aid wound healing an epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. The present sequence

CC represents a novel human secreted protein of the invention
XX
SQ Sequence 70 AA;

Query Match 100.0%; Score 260; DB 5; Length 70;
Best local Similarity 100.0%; Pred. No. 4.3e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMGMLTKIKMCAMWAVYCSFISFANSRSSEDTKQMSSEFM 49
Db 1 MNLIGMIFSMGMLTKIKMCAMWAVYCSFISFANSRSSEDTKQMSSEFM 49

RESULT 13
ABB72197
ID ABB72197 standard; protein: 70 AA.

AC ABB72197;
XX
DT 04-APR-2002 (first entry)

XX Human protein isolated from skin cells SEQ ID NO: 313.

XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
XX developmental defect; inflammatory disease; dermatological; vulnary;
XX immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX Homo sapiens.

XX WO200190357-A1.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-NZ000099.

XX 24-MAY-2000; 2000US-0206650P.

XX 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Sleeman M, Onrust R, Morrison JG;
PI Kumble KD;

XX WPI; 2002-122020/16.

XX New polynucleotides and polypeptides encoded by the polynucleotides
XX isolated from skin cells, useful for treating skin wounds, cancers,
XX growth and developmental defects, inflammatory diseases, or for
XX modulating immune responses.

XX Example 2; Page 207; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs
XX isolated from human, murine and rat skin cell libraries. The sequences
XX can be used in the development of therapeutic agents useful in the
XX treatment of skin diseases, including skin wounds, cancer, growth
XX defects, developmental defects and inflammatory diseases. The proteins
XX have important roles in the induction of hair growth, cell proliferation
XX and cell-cell interaction, in maintaining tissue integrity, in wound
XX healing and in modulating immune responses. The present sequence is a
XX polypeptide of the invention

XX Sequence 70 AA;

Query Match 100.0%; Score 260; DB 5; Length 70;
Best local Similarity 100.0%; Pred. No. 4.3e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLIGMIFSMGMLTKIKMCAMWAVYCSFISFANSRSSEDTKQMSSEFM 49
Db 1 MNLIGMIFSMGMLTKIKMCAMWAVYCSFISFANSRSSEDTKQMSSEFM 49

RESULT 14
AA87077
ID AA87077 standard; protein: 71 AA.

AC AA87077;

DT 09-MAY-2000 (first entry)

XX Human secreted protein sequence SEQ ID NO:116.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX antiinflammatory; neurotrophic; neuroprotective; antiallergic; cancer;
XX tumor; neurodegenerative disorder; developmental abnormality; allergy;
XX foetal deficiency; blood disorder; immune system disorder; arthritis;
XX autoimmune disease; hepatic disease; renal disease; inflammation;
XX Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
XX infection; AIDS; spinal cord injury; transplant rejection; diabetes;
XX asthma; sepsis; acne; psoriasis; cardiovascular disorder;
XX reproductive disorder; gastrointestinal disorder; respiratory disorder;
XX metabolic disorder; food additive; preservative; chromosome 19.

XX Homo sapiens.

XX WO200004140-A1.

XX 27-JAN-2000.

XX 14-JUL-1999; 99WO-US015849.

XX 15-JUL-1998; 98US-0092921P.

XX 15-JUL-1998; 98US-0092922P.

XX 15-JUL-1998; 98US-0092956P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
PI Lafleur DW, Edner R, Olsen HS, Brewer LA, Florence KA, Young PE;
PI Mucenski M, Endress GA, Soppet DR;

XX WPI; 2000-161128/14.

XX N-PDB; AA298030.

XX New isolated human genes, useful for diagnosis and treatment of, e.g.
XX cancers, neurological or blood disorders.

XX Claim 11; Page 380; 494pp; English.

XX The polynucleotide sequences given in AA298017 to AA298108 encode the
XX human secreted proteins given in AA87064 to AA87223. Human secreted
XX protein can have activities based on the tissues and cells the genes are
XX expressed in. Examples of activities include: cytostatic;
XX immunosuppressive; antiinflammatory; neurotrophic; neuroprotective; and
XX antiallergic. The polynucleotides and their corresponding secreted
XX polypeptides are useful for preventing, treating or ameliorating medical
XX conditions, e.g. by protein or gene therapy. Also pathological conditions
XX can be diagnosed by determining the amount of the new polypeptides in a
XX sample or by determining the presence of mutations in the new
XX polynucleotides. Human secreted protein 8 and their polynucleotides can
XX be used for developing products for the diagnosis or treatment of cancer,
XX tumours, neurodegenerative disorders, developmental abnormalities and
XX foetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, hepatic and renal diseases, inflammation, allergies,
XX Alzheimer's disease, behavioural disorders, schizophrenia, osteoporosis,
XX arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
XX diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
XX reproductive disorders, gastrointestinal disorders, respiratory disorders
XX and metabolic disorders. The proteins or polynucleotides can also be used
XX as food additives or preservatives. The proteins are also useful for
XX identifying their binding partners. AA298008 to AA298016 and AA87063 are
XX sequence used in the exemplification of the present invention

XX Sequence 71 AA;

Query Match 100.0%; Score 260; DB 3; Length 71;
Best Local Similarity 100.0%; Pred. No. 4,4e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MNLIGMIFSMCGMLKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49
1 MNLIGMIFSMCGMLKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49
Db

RESULT 15

AAV59659

ID AAV59659 standard; protein; 106 AA.

XX AAV59659;

XX 18-JAN-2000 (first entry)

XX Secreted protein 108-003-5-0-ES-FL.

XX Secreted protein; fingerprint identification technique;

XX chromosome mapping; human; hereditary disease; diagnosis; cancer;

XX hyperlipidaemia; cardiovascular; neurodegenerative disorder; therapy;

XX autoimmune disease; rheumatic disease; embryonic disorder; myopathy;

XX renal injury; amino aciduria; hypoglycaemia; male rat infertility;

XX hypertension.

XX Homo sapiens.

XX MO940189-A2.

XX 12-AUG-1999.

XX 09-FEB-1999; 99MO-IB000282.

XX 09-FEB-1998; 98US-0074121P.

XX 13-APR-1998; 98US-0081563P.

XX 10-AUG-1998; 98US-0096116P.

XX 04-SEP-1998; 98US-0099273P.

XX (BEST) GENSET.

XX Bougueleret L, Duclert A, Dumas Milne Edwards J;

XX WPI: 1999-600966/51.

XX N-PSDB; AA240787.

XX Extended cDNAs useful for expressing secreted proteins and to obtain

XX specific antibodies.

XX Claim 10; Page 191; 244pp; English.

XX This sequence represents a human secreted protein of the invention. The

XX extended cDNAs (or genomic DNAs obtainable from them) may be used to

XX prepare PCR primers and probes. These are useful for forensic matching or

XX positive identification by DNA sequencing. They may also be used in

XX alternative fingerprint identification techniques. Antibodies against the

XX proteins encoded by the extended cDNAs are useful in identification of

XX tissue types or cell species, as well as identifying tissue specific

XX soluble proteins. The sequences can be used for chromosome mapping and

XX identification of genes associated with hereditary diseases or drug

XX response. Signal sequences from the cDNAs can be used in construction of

XX secretion vectors. Other sequences derived from the extended cDNAs can be

XX used to clone upstream genomic DNA sequences including promoters. This is

XX in turn useful for identifying proteins that interact with promoter

XX sequences. Some of the proteins may be useful in diagnosing and treating

XX several disorders including, but not limited to: cancer, hyperlipidaemia,

XX cardiovascular and neurodegenerative disorders, autoimmune diseases, and

XX rheumatic diseases, embryonic disorders, hypertension, renal injury,

XX amino acidurias, hypoglycaemia, male rat infertility and myopathies

XX Sequence 106 AA;

XX SQ

XX Query Match 100.0%; Score 260; DB 2; Length 106;

Best Local Similarity 100.0%; Pred. No. 6,9e-30;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNLIGMIFSMCGMLKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 49

Db 37 MNLIGMIFSMCGMLKLKWCAMVAVYCSFISFANSRSSEDTKQMSSEFM 85

Search completed: March 6, 2005, 08:06:00
Job time : 126 secs

